

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
16 May 2002 (16.05.2002)

PCT

(10) International Publication Number  
**WO 02/38580 A1**

(51) International Patent Classification<sup>7</sup>: **C07H 21/02**,  
21/04, C07K 1/00, 5/00, C12N 15/00, 15/63, C12Q 1/00,  
G01N 33/53

**KEEFE, Anthony, D.**; 97H Thorndike Street, Apt. #1,  
Cambridge, MA 02141 (US).

(21) International Application Number: PCT/US00/41717

(74) Agent: **ELBING, Karen, L.**; Clark & Elbing LLP, 176  
Federal Street, Boston, MA 02114-2214 (US).

(22) International Filing Date: 31 October 2000 (31.10.2000)

(81) Designated States (*national*): CA, JP.

(25) Filing Language: English

(84) Designated States (*regional*): European patent (AT, BE,  
CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,  
NL, PT, SE).

(26) Publication Language: English

(71) Applicant: **THE GENERAL HOSPITAL CORPORA-**  
**TION** [US/US]; 55 Fruit Street, Boston, MA 02114 (US).

Published:  
— with international search report

(72) Inventors: **SZOSTAK, Jack, W.**; 390 Marlborough  
Street, Boston, MA 02115 (US). **WILSON, David, S.,**  
**Jr.**; 24702 Broadmore Avenue, Hayward, CA 94544 (US).

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.



**WO 02/38580 A1**

(54) Title: STREPTAVIDIN-BINDING PEPTIDES AND USES THEREOF

(57) Abstract: The invention provides peptides with high affinity for streptavidin. These peptides may be expressed as part of fusion proteins to facilitate the detection, quantitation, and purification of proteins of interest.

STREPTAVIDIN-BINDING PEPTIDES AND USES THEREOF

5

Statement as to Federally Sponsored Research

This invention was funded by grant number R01GM53936 from the National Institutes of Health and grant number NCC-2-1069 from NASA. The government may have certain rights in the invention.

10

Background of the Invention

In general, the invention features novel compounds and methods for purifying or detecting proteins of interest.

Determining the enzymatic activity, binding specificity, or three-dimensional structure of a protein often requires the purification of the protein from a complex mixture of other components, such as compounds present in a cell lysate or *in vitro* translation extract. With the number of known proteins increasing dramatically as a result of whole genome sequencing projects, it has become crucial to find alternatives to traditional, time-consuming monoclonal antibody production for generating affinity reagents for the detection and purification of proteins. In addition, purifying a novel protein using traditional column chromatography methods often requires much trial and error to develop a purification protocol that results in the recovery of the protein in high yield and purity.

Thus, purification methods are needed that may be generally applied to proteins of interest, that utilize inexpensive reagents, and that result in highly purified protein without requiring multiple chromatography steps.

### Summary of the Invention

The purpose of the present invention is to provide improved reagents for the purification, detection, or quantitation of proteins of interest. In particular, the high affinity, streptavidin-binding peptides of the present invention may be used  
5 as affinity tags for the purification of fusion proteins containing proteins of interest.

Accordingly, in a first aspect, the invention provides a peptide which binds streptavidin with a dissociation constant less than  $10\ \mu\text{M}$  (that is, binds streptavidin more tightly than a  $K_d$  of  $10\ \mu\text{M}$ ) and which is not disulfide bonded  
10 or cyclized. Preferably, the dissociation constant is equal to or less than  $5\ \mu\text{M}$ ,  $1\ \mu\text{M}$ ,  $100\ \text{nM}$ ,  $50\ \text{nM}$ ,  $25\ \text{nM}$ ,  $10\ \text{nM}$ , or even  $5\ \text{nM}$ . In one preferred embodiment, the dissociation constant is less than  $10\ \mu\text{M}$ ,  $5\ \mu\text{M}$ ,  $1\ \mu\text{M}$ ,  $100\ \text{nM}$ ,  $50\ \text{nM}$ , or  $25\ \text{nM}$ ; and greater than  $0.01\ \text{nM}$ ,  $0.1\ \text{nM}$ ,  $1\ \text{nM}$ ,  $5\ \text{nM}$ , or  $10\ \text{nM}$ . In another preferred embodiment, the value of the dissociation constant is contained  
15 in one of the following ranges:  $5\ \mu\text{M}$  to  $1\ \mu\text{M}$ ,  $1\ \mu\text{M}$  to  $100\ \text{nM}$ ,  $100\ \text{nM}$  to  $50\ \text{nM}$ ,  $50\ \text{nM}$  to  $25\ \text{nM}$ ,  $25\ \text{nM}$  to  $10\ \text{nM}$ ,  $10\ \text{nM}$  to  $5\ \text{nM}$ ,  $5\ \text{nM}$  to  $1\ \text{nM}$ , or  $5\ \text{nM}$  to  $0.1\ \text{nM}$ , inclusive.

In a related aspect, the invention provides a peptide which binds streptavidin with a dissociation constant less than  $10\ \mu\text{M}$ . The amino acid  
20 sequence of the peptide does not contain an HPQ, HPM, HPN, or HQP motif. Preferably, the dissociation constant is equal to or less than  $5\ \mu\text{M}$ ,  $1\ \mu\text{M}$ ,  $100\ \text{nM}$ ,  $50\ \text{nM}$ ,  $25\ \text{nM}$ ,  $10\ \text{nM}$ , or  $5\ \text{nM}$ . In one preferred embodiment, the dissociation constant is less than  $10\ \mu\text{M}$ ,  $5\ \mu\text{M}$ ,  $1\ \mu\text{M}$ ,  $100\ \text{nM}$ ,  $50\ \text{nM}$ , or  $25\ \text{nM}$ ; and greater than  $0.01\ \text{nM}$ ,  $0.1\ \text{nM}$ ,  $1\ \text{nM}$ ,  $5\ \text{nM}$ , or  $10\ \text{nM}$ . In another preferred embodiment,  
25 the value of the dissociation constant is contained in one of the following ranges:  $5\ \mu\text{M}$  to  $1\ \mu\text{M}$ ,  $1\ \mu\text{M}$  to  $100\ \text{nM}$ ,  $100\ \text{nM}$  to  $50\ \text{nM}$ ,  $50\ \text{nM}$  to  $25\ \text{nM}$ ,  $25\ \text{nM}$  to  $10\ \text{nM}$ ,  $10\ \text{nM}$  to  $5\ \text{nM}$ ,  $5\ \text{nM}$  to  $1\ \text{nM}$ , or  $5\ \text{nM}$  to  $0.1\ \text{nM}$ , inclusive.

In another related aspect, the invention provides a peptide which binds streptavidin with a dissociation constant less than  $23\ \text{nM}$ ,  $10\ \text{nM}$ , or  
30  $5\ \text{nM}$ . In one preferred embodiment, the peptide is disulfide bonded or cyclized.

In another preferred embodiment, the dissociation constant is less than 23 nM, 10 nM, or 5 nM; and greater than 0.01 nM, 0.1 nM, or 1 nM. In another preferred embodiment, the value of the dissociation constant is contained in one of the following ranges: 20 nM to 10 nM, 10 nM to 5 nM, 5 nM to 1 nM, or 5 nM to 0.1 nM, inclusive.

In other related aspects, the invention provides nucleic acids encoding the peptides of the present invention, and vectors that include such nucleic acids.

In addition, standard gene fusion techniques may be used to generate fusion nucleic acids that encode fusion proteins which include a peptide of the present invention and a protein of interest. The fusion proteins may be purified, detected, or quantified based on the high affinity of the peptides for streptavidin.

Accordingly, in one such aspect, the invention provides a fusion protein including a protein of interest covalently linked to one of the following peptides:

(a) a peptide which binds streptavidin with a dissociation constant less than 10  $\mu$ M and which is not disulfide bonded or cyclized, (b) a peptide which binds streptavidin with a dissociation constant less than 10  $\mu$ M and which does not contain an HPQ, HPM, HPN, or HQP motif, or (c) a peptide which binds streptavidin with a dissociation constant less than 23 nM. In preferred embodiments, the peptide is attached to the amino-terminus or the carboxy-terminus of the protein of interest, or the peptide is positioned between the amino and carboxy-termini of the protein of interest. Preferably, the peptide is linked to the protein of interest by a linker which includes a protease-sensitive site.

In related aspects, the invention provides nucleic acids encoding the fusion proteins of the present invention, and vectors that include these fusion nucleic acids.

In addition, the invention provides a method of producing a fusion protein of the present invention. This method includes transfecting a vector having a nucleic acid sequence encoding the fusion protein into a suitable host cell and culturing the host cell under conditions appropriate for expression of the fusion protein.

The fusion proteins described herein may be used in methods for purifying proteins of interest from samples. Such a method involves expressing the protein of interest as a fusion protein covalently linked to one of the following peptides: (a) a peptide which binds streptavidin with a dissociation constant less than 10  $\mu$ M and which is not disulfide bonded or cyclized, (b) a peptide which binds streptavidin with a dissociation constant less than 10  $\mu$ M and which does not contain an HPQ, HPM, HPN, or HQP motif, or (c) a peptide which binds streptavidin with a dissociation constant less than 23 nM. A sample containing the fusion protein is contacted with streptavidin under conditions that allow complex formation between the fusion protein and streptavidin. The complex is isolated, and the fusion protein is recovered from the complex, thereby purifying the protein of interest from the sample. In one preferred embodiment, the protein of interest is recovered from the fusion protein by cleaving the streptavidin-binding peptide from the fusion protein.

In yet another aspect, the invention provides a method of detecting the presence of a fusion protein of the present invention in a sample. This method includes (a) contacting the sample with streptavidin under conditions that allow complex formation between the fusion protein and streptavidin, (b) isolating the complex, and (c) detecting the presence of streptavidin in the complex or following recovery from the complex. The presence of streptavidin indicates the presence of the fusion protein in the sample. Preferably, step (c) also involves measuring the amount of streptavidin in the complex or following recovery from the complex. The amount of fusion protein in the sample is correlated with, and may be calculated from, the measured amount of streptavidin. For example, for a fusion protein containing a peptide that binds one molecule of streptavidin per molecule of peptide, the amount of fusion protein in the sample is predicted to be approximately the same as the amount of streptavidin measured. In one preferred embodiment, the amount of streptavidin is determined using Western or ELISA analysis with an antibody that reacts with streptavidin or that reacts with a compound that is covalently linked to streptavidin. In another preferred

embodiment, streptavidin is covalently linked to an enzyme, radiolabel, fluorescent label, or other detectable group, and the amount of streptavidin is determined using standard techniques based on a characteristic of the detectable group such as its enzyme activity, radioactivity, or fluorescence.

5 In preferred embodiments of various aspects of the invention, the amino acid sequence of the peptide includes at least 10, 25, 50, 75, or 100 consecutive amino acids or consists of between 5 and 150, 10 and 100, 20 and 75, or 30 and 50 amino acids, inclusive, of any one of SEQ ID Nos. 1-29 or 35. Preferably, the amino acid sequence of the peptides includes an LPQ, QPQ, EPQ, HPA, HPD, or  
10 HPL motif. In other preferred embodiments, the amino acid sequence includes any one of SEQ ID Nos. 1-29 or 35. In still other preferred embodiments, the peptide has an amino acid sequence that is at least 20, 30, 40, 50, 60, 70, 80, 90, 95, or 100% identical to any one of SEQ ID Nos. 1-29 or 35.

It is also contemplated that the affinity of the peptides of the present  
15 invention for streptavidin may be increased by incorporating disulfide bonds into, or cyclizing, the peptides. By constraining the peptides, the amount of disorder inherent in the peptides (i.e., entropy) decreases, and thus binding of these peptides to streptavidin may require less energy. It is also contemplated that the three-dimensional structure of peptides of the invention bound to streptavidin may  
20 be experimentally determined or modeled based on the known crystal structure of streptavidin and used to determine possible modifications to the peptides that may further improve their affinity for streptavidin.

As used herein, by "nucleic acid" is meant a sequence of two or more covalently bonded naturally-occurring or modified deoxyribonucleotides or  
25 ribonucleotides.

By "peptide" is meant a sequence of two or more covalently bonded naturally-occurring or modified amino acids. The terms "peptide" and "protein" are used interchangeably herein.

By "covalently linked" is meant covalently bonded or connected through a series of covalent bonds. A group that is covalently linked to a protein may be attached to the amino-terminus, carboxy-terminus, between the amino- and carboxy-termini, or to a side chain of an amino acid in the protein.

5 By "streptavidin" is meant any streptavidin molecule or fragment thereof or any protein that has an amino acid sequence that is at least 80, 90, 95, or 100% identical to a streptavidin molecule or fragment thereof (see, for example, Hauptle *et al.* J. Biol. Chem. 258: 305, 1983). A preferred fragment of streptavidin is "core" streptavidin, which is a proteolytic cleavage product of  
10 streptavidin (Bayer *et al.* Biochem. J. 259,369-376, 1989). Preferably, a streptavidin molecule or fragment thereof is capable of binding biotin or any other streptavidin-binding molecule. Streptavidin or a streptavidin fragment may be modified chemically or through gene fusion technology or protein synthesis so that it is covalently linked to an enzyme, radiolabel, fluorescent label, or other  
15 detectable group. These detectable groups may be used to determine the presence or location of a streptavidin-bound fusion protein in a cell or sample or to quantify the amount of a streptavidin-bound fusion protein, using standard methods.

By "cyclized" is meant nonlinear. A peptide may be cyclized by the  
20 formation of a covalent bond between the N-terminal amino group of the peptide or the side-chain of a residue and the C-terminal carboxyl group or the side-chain of a residue. For example, a peptide lactam may be formed by the cyclization between the N-terminal amino group or an amino group of an amino acid side-chain and the C-terminal carboxyl group or a carboxyl or amide containing side-  
25 chain. Other possible cyclizations include the formation of a thioether by the reaction of a thiol group in a cysteine side-chain with the N-terminal amino group, C-terminal carboxyl group, or the side-chain of another amino acid. A disulfide bond may also be formed between two cysteine residues. As used herein, a "non-cyclized peptide" is a linear peptide that does not have any of the above  
30 cyclizations.

By "dissociation constant" is meant the dissociation constant for binding streptavidin as measured using the electrophoretic mobility shift assay described herein. By "less than" a particular dissociation constant is meant capable of binding streptavidin more tightly than the strength of binding represented by a particular dissociation constant.

By "purifying" is meant separating a compound, for example, a protein, from other components that naturally accompany it. Typically, a protein is substantially pure when it is at least 50%, by weight, free from proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the protein is at least 75%, more preferably, at least 90%, and most preferably, at least 99%, by weight, pure. In other preferred embodiments, the protein is at least 2, 5, 10, 25, 50, or 100 times as pure as the starting material. Purity may be assayed by any appropriate method, such as polyacrylamide gel electrophoresis, column chromatography, optical density, HPLC analysis, western analysis, or ELISA (see, for example, Ausubel *et al.*, Current Protocols in Molecular Biology, John Wiley & Sons, New York, 2000).

By "recovered from the complex" is meant physically separated from the complex of streptavidin and the fusion protein. For example the streptavidin-bound fusion protein may be incubated under conditions that reduce the affinity of the fusion protein for streptavidin (*i.e.*, at low or high salt concentrations or at low or high pH values) or incubated in the presence of molecules that compete with the fusion protein for binding streptavidin. Preferably, either the fusion protein or the streptavidin that has been released from the complex is isolated using standard procedures, such as column chromatography, polyacrylamide gel electrophoresis, HPLC, or western analysis.

The present invention provides a number of advantages related to the detection and purification of proteins of interest. For example, because the present methods do not require the generation of an antibody or other affinity reagent that is specific for each protein of interest, these methods may be universally applied to any protein. In addition, if desired, the streptavidin-binding



peptide may be connected to the protein of interest through a protease cleavable linker, allowing removal of the peptide after purification of the fusion protein. Using the methods described herein, purification of a fusion protein based on its affinity for streptavidin has allowed the isolation of the fusion protein in

5 significantly higher purity than that obtained using a hexahistidine affinity tag or maltose-binding protein affinity tag. Moreover, streptavidin is an inexpensive reagent that may be purchased unmodified or covalently labeled with a detectable group (such as FITC-streptavidin or alkaline phosphatase-conjugated streptavidin) or with a chromatography matrix (such as streptavidin-agarose). The availability

10 of these reagents simplifies methods for detecting and purifying the fusion proteins of the present invention.

Other features and advantages of the invention will be apparent from the following detailed description and from the claims.

15 Brief Description of the Drawings

Figure 1A is a schematic illustration of an *in vitro* selection process according to the invention, showing the structure of the library and the selection scheme. Members of the DNA library have, from the 5' to 3' end, a T7 RNA polymerase promoter (T7), a tobacco mosaic virus translation enhancer (TMV), a

20 start codon (ATG), 88 random amino acids, a hexahistidine tag (H6), and a 3' constant region (Const).

Figure 1B is a picture of an SDS-PAGE gel of samples from the library at different stages of preparation. The first lane shows the result of translating the mRNA display template with <sup>35</sup>S-methionine. Most of the counts represent free peptide (free pep), but a significant amount of mRNA-peptide covalent fusions are

25 also present (mRNA-pep). There is also another band that is independent of added template (NS, non-specific), and some counts remain in the gel well. The band corresponding to the mRNA-peptide can be shifted to a position slightly higher than that for the free peptide by the addition of RNase A. The remaining lanes

show the result of successive oligo-dT and Ni-NTA purifications, and finally reverse transcription (RT).

Figure 2A is a bar graph showing the fraction of  $^{35}\text{S}$  counts from the displayed peptides that bound to streptavidin and eluted with biotin, at each round of selection. Figure 2B is a graph showing the elution profile for the peptide library generated from the output of the seventh round of selection in Figure 2A. The first fraction represents the flow-through. Biotin was added at the point indicated. The plot compares the binding of the intact, reverse-transcribed, displayed peptides (mRNA-pep), the same sample treated with RNase A, and the RNase-treated sample applied to a streptavidin column pre-saturated with biotin (excess biotin was washed away prior to exposing the library to the matrix).

Figure 3 is a list of the sequences of 20 clones from the seventh round of selection (SEQ ID Nos.: 1-20). The "#" column indicates the number of times each sequence was observed. The HPQ sequence is in bold type. Non-random sequences at the termini are underlined. The six C-terminal-most residues are not shown.

Figure 4A is a picture of a native gel showing an electrophoretic mobility shift (EMSA) analysis demonstrating the binding of four different DNA-tagged peptides to streptavidin. The migration of each clone is shown in the absence (-) and presence (+) of 1  $\mu\text{M}$  streptavidin. Some of the clones show multiple bands, presumably representing different conformations. The arrows show the position of the gel well, which often contains a fraction of the counts. Figure 4B is a picture of a native gel showing the titration of the full-length clone SB19 with streptavidin. The streptavidin concentration in each lane, from left to right, is: 3.8, 6.6, 10, 15, 23, 35, and 61 nM. Figure 4C is a curve fit of the data shown in Figure 4B (the fraction of peptide bound could not be accurately determined for the point with the lowest concentration of streptavidin). Assuming that the peptide is homogeneous and 100% active, the data from this experiment give a  $K_d$  of 10 nM for the binding of peptide SB19 to streptavidin.

Figure 5 is a list of the sequences of truncation mutants of peptide SB19 (SEQ ID Nos.: 21-29). The full-length (FL), C-terminal deleted (C1-C4), N-terminal deleted (N1-N3), and point mutated (M1) peptide sequences are shown. The "% binding" refers to the performance of these peptides in the streptavidin column-binding assay.

Figure 6A is the nucleotide sequence of the plasmid used for expression of a fusion protein containing a streptavidin-binding peptide (SEQ ID No.: 37). Figure 6B is the amino acid sequence of the encoded protein (SEQ ID No.: 38) which contains, from the amino- to carboxy-terminus, maltose-binding protein, a streptavidin-binding peptide (SEQ ID No.: 35, Fig. 7A), a hexahistidine tag, and another peptide called 2r18-19dN. Figure 6C is the amino acid sequence of 2r18-19dN (SEQ ID No.: 39).

Figure 7A is the amino acid sequence of the streptavidin-binding peptide (SEQ ID No.: 35) used as an affinity tag for the purification of the fusion protein listed in Fig. 6B. This peptide contains the first 38 amino acids of the SB19-C4 peptide (Fig. 5). Figure 7B is a picture of an SDS-PAGE gel showing the purity of the fusion protein after elution from the streptavidin column (lane 2) compared to the purity of the *E. coli* lysate that was applied to the column (lane 1).

Figures 8A-8F are schematic illustrations of the pre-selection method. Figure 8A is an illustration of an mRNA display template terminating in puromycin in which the tobacco mosaic virus translation enhancer sequence (TMV), the initiating methionine codon (AUG), and the sections of the open reading frame encoding the two protein affinity tags (FLAG and His<sub>6</sub>) are labeled. Figure 8B is an illustration of an mRNA display template that is free of frameshifts and premature stop codons and thus encodes a full-length protein containing both affinity tags. Figure 8C illustrates an mRNA display template that has initiated internally and displays the corresponding truncated protein lacking the N-terminal FLAG tag. Figure 8D shows an mRNA display template that has a deletion in its open reading frame and thus displays the corresponding frameshifted protein lacking the C-terminal His<sub>6</sub> tag. Figure 8E illustrates the

reverse transcription of the mRNA display template from Fig. 8B that was purified based on the presence of both protein affinity tags in the encoded protein. Figure 8F shows the cleavage sites for Type IIS restriction enzymes which are encoded in each cassette. Ligation of pre-selected cassettes which have been  
5 cleaved with these enzymes yields the full-length DNA library.

Figure 9A is the polynucleotide sequence of the vector encoding a fusion protein containing maltose-binding protein, a streptavidin-binding peptide (SEQ ID No.: 35, Fig. 7A), and a hexahistidine tag. Figure 9B is the amino acid sequence of the encoded fusion protein. The sequence of the streptavidin-binding  
10 peptide which contains the first 38 amino acids of the SB19-C4 peptide is underlined.

Figure 10A is a graph of the biacore response units over various lengths of time for the dissociation of streptavidin from the fusion protein listed in Fig. 9B immobilized on a biacore chip. For line "a," the streptavidin concentration is 23  
15  $\mu\text{M}$ ; for line "b," the concentration is 11.5  $\mu\text{M}$ , and for line "c," the concentration is 5.75  $\mu\text{M}$ . This data was used to calculate an upper limit of  $2 \times 10^{-3}/\text{s}$  for the dissociation rate,  $k_d$ . Figure 10B is a graph showing the association and subsequent dissociation of streptavidin from the immobilized fusion protein. For lines "a" through "f," the streptavidin concentrations are 1.6, 0.8, 0.4, 0.2, 0.1,  
20 and 0.05  $\mu\text{M}$ , respectively. This data was used to calculate an association rate,  $k_a$ , of  $5 \times 10^4/\text{M}/\text{s}$ .

#### Detailed Description

The present methods stem from the discovery of peptides that have  
25 unusually high affinities for streptavidin ( $K_d$  of less than 10  $\mu\text{M}$ ). These peptides were selected from a library of randomized, non-constrained peptides using the mRNA display method. The high affinity of the selected peptides was particularly surprising, given the fact that non-constrained linear peptide libraries generally do not yield high affinity ligands to proteins, except in cases where the  
30 protein normally functions in peptide recognition (Clackton *et al.*, Trends Biotech

12:173-184 (1994); Katz, *Annu. Rev. Biophys. Biomol. Struct.* 26:27-45, 1997). Many other peptides with high affinity for streptavidin may be isolated using the mRNA display method or any other selection method, such as ribosome display (Roberts, *Curr. Opin. Chem. Biol.* 3(3):268-73, 1999), or phage display (U.S.

5 Patent No. 5,821,047).

The binding characteristics of exemplary selected streptavidin-binding peptides are described in Table 1, and the sequences of these peptides are listed in Fig. 3. The first column of Table 1 lists the peptide name (SB1 - SB20). For comparison, a non-selected sequence with two HPQ motifs spaced by 19 residues  
10 (called "non-selected") is listed in row one. SB19-C4 is a truncation mutant of peptide SB19, described below. The peptides are grouped according to the number of HPQ and similar tripeptide motifs they possess. The second column shows the number of tripeptide motifs in each peptide, and the number of amino acid residues separating them. The third column represents the percentage of  
15 peptide binding and specifically eluting from a streptavidin column. This percentage ranged from 8.3% to as high as 88% for the selected peptides, compared to only 0.16% for the control, non-selected peptide with two HPQ motifs.

The fourth column shows the  $K_d$ , when known, for the interaction between  
20 streptavidin and the peptides, as measured in the EMSA assay described herein. The standard deviation in the  $K_d$  is shown in the fifth column, based on the number of independent measurements (n, shown in parentheses). The dissociation constant ranged from 110 nM for peptide SB5 to 4.8 nM for peptides SB2.

25

Table 1

	Peptide	Structure	% binding and eluting	$K_d$ (nM)	Standard deviation (n)
5	Non-selected HPQ 19 HPQ		0.16		
	<i>Two HPQ motifs</i>				
	SB1	HPQ 19 HPQ	86	50	5.7 (4)
10	SB2	HPQ 19 HPQ	48	4.8	0.91 (8)
	SB3	HPQ 23 HPQ	20		
	SB4	HPQ 43 HPQ	49		
	SB5	HPQ 52 HPQ	72	110	22 (6)
15	<i>One HPQ and one similar tripeptide motif</i>				
	SB6	HPL 4 HPQ	49		
	SB7	HPD 7 HPQ	28		
	SB8	HPQ 12 HPL	27		
	SB9	HPQ 12 HP	64		
20	SB10	HPQ 21 QPQ	15		
	SB11	HPQ 28 HPA	68		
	SB12	HPQ 30 EPQ	73		
	SB13	HPQ 32 EPQ	64		
	SB14	HPQ 43 HPL	11		
25	SB15	QPQ 50 HPQ	44	92	16 (4)
	SB16	HPQ 74 LPQ	50		
	<i>One HPQ motif</i>				
	SB17		8.3		
30	SB18		58		
	SB19		85	10	1.8 (10)
	SB19 - C4		88	4.9	0.88 (10)
	<i>No HPQ motif</i>				
35	SB20	HPL	34		

To further characterize the binding of the selected peptides to streptavidin, truncation mutants for peptide SB19 were constructed to determine which regions were necessary for high affinity streptavidin-binding (Fig. 5). Deletion of up to 56 residues had no observable effect on the binding strength. For example, peptide SB19-C4 retained only the first 38 residues from the selected construct (plus the C-terminal sequence MMSGGCKLG, SEQ ID No.: 36) and had a dissociation constant of 4.9 nM for streptavidin (Table 1). In contrast, N-terminal truncation mutations (N1-N3) resulted in a lower percentage of the encoded peptide specifically eluting from the streptavidin column (0.058 to 69% for the truncation mutants compared to 85% for full length SB19). These results suggested that the determinants for binding streptavidin were spread throughout the N-terminal 38 residues of the SB19 peptide.

High affinity streptavidin-binding peptides, such as those shown in Table 1, have a number of uses. For example, these peptides may be used for protein purification by expressing a protein of interest as a fusion protein joined to one or more of the streptavidin-binding peptides of the invention. In one such purification method, a sample containing the fusion protein is incubated with immobilized streptavidin. Proteins with no or weak affinity for streptavidin are washed away, and the fusion protein is then selectively eluted from the streptavidin matrix by addition of biotin, a biotin analog, another streptavidin-binding peptide, or any compound that competes with the fusion protein for binding to the matrix. Alternatively, the fusion protein may be eluted from the matrix by increasing or decreasing the pH of the buffer applied to the matrix.

As described in detail below, this general protocol was used in a one-step purification of a fusion protein containing a streptavidin-binding peptide from an *E. coli* extract, resulting in a high yield of very pure protein. This fusion protein contained the first 38 amino acids of the SB19-C4 peptide, which due to its small size was not expected to affect the three-dimensional structure or activity of the covalently-linked protein of interest. Purification of fusion proteins containing

other streptavidin-binding peptides of the present invention may be performed similarly.

In addition, various modifications of the above purification protocol would be apparent to one skilled in the art (see, for example, Ausubel *et al.*,  
5 *supra*), and such modifications are included in the invention. In particular, use of the streptavidin-binding peptides as affinity tags is desirable for high throughput protein production and purification. For example, purification of fusion proteins in a multi-well format may be conducted using magnetic streptavidin beads that are washed and eluted robotically. The methods of the present invention may  
10 also be adapted to purify fusion proteins from *in vitro* translation mixtures or from other extracts, such as those from prokaryotic, yeast, insect, or mammalian cells, using standard techniques. If necessary, avidin may be added to the extract to bind any free biotin in the extract before contacting a sample from the extract with streptavidin. Allowing any free biotin to bind avidin may prevent biotin  
15 from competing with the streptavidin-binding peptides for binding to streptavidin.

If desired, the presence of a fusion protein of the invention in a sample may be detected by incubating the fusion protein with streptavidin (*i.e.*, unlabeled streptavidin or streptavidin that is labeled with a detectable group) under  
20 conditions that allow streptavidin to bind the fusion protein. Preferably, the unbound streptavidin is separated from the streptavidin-bound fusion protein. Then, the streptavidin that is bound to the fusion protein is detected. Alternatively, the streptavidin bound to the fusion protein is physically separated from the fusion protein and then detected, using standard methods. For example,  
25 to detect streptavidin that is bound to the fusion protein or that has been separated from the fusion protein, Western or ELISA analysis may be performed using an antibody that reacts with streptavidin or that reacts with a compound that is covalently linked to streptavidin. If streptavidin is covalently linked to an enzyme, radiolabel, fluorescent label, or other detectable group, the amount of  
30 streptavidin may be determined using standard techniques based on a



characteristic of the detectable group such as its enzyme activity, radioactivity, or fluorescence (see, for example, Ausubel *et al.*, *supra*). Alternatively, streptavidin may be contacted with a streptavidin-binding compound that is covalently linked to an enzyme, radiolabel, fluorescent label, or other detectable group, and the  
5 detectable group may be assayed as described herein.

We have also developed an improved method to generate synthetic DNA libraries encoding full-length proteins, which may be used in a variety of selection methods to isolate proteins with desired binding affinities or activities. The generation of libraries of proteins containing a desired number of amino  
10 acids is often limited by the number of internal initiation events that result in truncated proteins and the number of frameshifts that result in either premature stop codons or the removal of desired stop codons. For example, during solid phase DNA synthesis, insertions and deletions which cause frameshifts may occur due to imperfect coupling and capping efficiencies. In addition, the  
15 random regions in DNA templates may encode stop codons, resulting in premature truncation of the encoded protein. To address these problems, we have developed a method in which small DNA cassettes are synthesized, and an *in vitro* selection using the mRNA display technology is performed to enrich the library of DNA cassettes for sequences encoding two protein affinity tags. These  
20 DNA cassettes lack frameshifts and premature stop codons. The selected DNA cassettes are then cleaved with restriction enzymes and ligated to generate the full-length DNA library (Figs. 8A-8F) (Cho *et al.*, J. Mol. Biol. 297:309-319, 2000).

In one preferred embodiment of this method, mRNA display templates that  
25 contain a translation enhancer sequence operably-linked to an open reading frame and that terminate in puromycin are generated as described previously (Cho *et al.*, *supra*). The open reading frame encodes two different protein affinity tags, such as a FLAG tag and a hexahistidine tag. Preferably, one of the tags is located at the amino-terminus of the encoded peptide, and the other tag is located at the  
30 carboxy-terminus. The mRNA display templates are *in vitro* translated to

generate mRNA displayed peptides (Cho *et al.*, supra). mRNA displayed peptides encoded by templates that do not contain frameshifts or premature stop codons should contain both affinity tags. In contrast, templates that contain frameshifts or premature stop codons encode peptides without the C-terminal affinity tag (Fig. 8D). Additionally, mRNA display templates that initiate internally produce peptides without the N-terminal affinity tag (Fig. 8C). The library of mRNA displayed peptides is enriched for peptides containing both affinity tags by purification of the mRNA displayed peptides based on the presence of these tags (see, for example, Ausubel *et al.*, supra). For example, the mRNA displayed peptides may be applied to a matrix designed to bind peptides containing one of the affinity tags, and the mRNA display peptides without the affinity tag are washed away. The mRNA display peptides containing the affinity tag are then eluted and applied to a second matrix designed to bind the other affinity tag. The mRNA display peptides recovered from this purification step are enriched for members containing both affinity tags and thus for full-length peptides. These mRNA displayed peptides are reversed transcribed to generate double-stranded DNA. The amplified DNA is then cleaved by restriction enzymes. Preferably, this restriction digestion removes the sequences encoding the affinity tags from the DNA cassettes. The cleaved DNA cassettes are then ligated to generate the full-length DNA templates.

The experiments described above were carried out as follows.

#### Generation of a Streptavidin-Binding Peptide Library

The mRNA display method for selecting peptides or proteins of interest takes advantage of the translation-terminating antibiotic puromycin, which functions by entering the A site of ribosomes and forming a covalent bond with the nascent peptide. By covalently attaching puromycin to the 3' end of an mRNA, a covalent link between a polypeptide and its encoding message can be achieved *in situ* during *in vitro* translation (Roberts *et al.*, Curr. Opin. Struct. Biol. 9:521-529, 1999; Liu *et al.*, Methods Enzymol. 318:268-293, 2000). These

mRNA-peptide fusions can then be purified and subjected to *in vitro* selection, yielding the isolation of novel peptide ligands.

A DNA library encoding polypeptides of 108 amino acids was synthesized as described (Cho *et al.*, *supra*). The library consisted of short cassettes  
5 concatamerized together. Each cassette encoded a random peptide with a pattern of polar *versus* non-polar amino acid side chains compatible with forming an amphipathic  $\alpha$ -helix or  $\beta$ -strand (Cho *et al.*, *supra*). The random region was 88 amino acids long, followed by a C-terminal invariant region containing a hexahistidine tag (Fig. 1A).

10 The library had a complexity of  $2.4 \times 10^{14}$  at the DNA level. It was transcribed using T7 RNA polymerase (Fig. 1A), after which a "linker" oligonucleotide was added to the 3' end using T4 DNA ligase as described (Liu *et al.*, *supra*; Cho *et al.*, *supra*). The linker consisted of a 21 nucleotide long dA stretch, followed by a polyethylene glycol linker, followed by the sequence dA-  
15 dC-dC-puromycin (Liu *et al.*, *supra*).

This puromycin-terminated mRNA was translated *in vitro*, using the Ambion (Austin, TX) *in vitro* translation kit under standard conditions for capped mRNA. The 10 mL reaction mixture was supplemented with 2 mCi  $^{35}\text{S}$ -methionine and a total methionine concentration of 10  $\mu\text{M}$ . The reaction mixture  
20 also included 300 nM of the library of puromycin-linked mRNA molecules. After 1 hour at 30°C,  $\text{MgCl}_2$  and KCl were added to 20 and 710 mM, respectively, and the reaction mixture was further incubated at room temperature for five minutes to increase the yield of displayed peptides. This *in vitro* translation produced  $1.2 \times 10^{14}$  polypeptides linked *via* the puromycin moiety to  
25 their encoding mRNAs.

These mRNA displayed peptides were then purified on oligo-dT cellulose (which binds to the oligo-dA sequence in the linker) to remove polypeptides not fused to mRNA. For this purification procedure, the reaction mixture was diluted 10-fold into oligo-dT-binding buffer (1M NaCl, 50 mM HEPES, 10 mM EDTA,

0.25% Triton X-100, and 5 mM 2-mercaptoethanol at pH 7.9) and 80 mg oligo-dT cellulose (type 7, Amersham-Pharmacia, Piscataway, NJ) and incubated with agitation at 4°C for 30 minutes. The mixture was applied to a column (Poly-Prep chromatography column, Biorad, Hercules, CA), drained, washed with 10 mL  
5 oligo-dT-binding buffer, washed with 10 mL oligo-dT-wash buffer (300 mM NaCl, 20 mM HEPES, 1 mM EDTA, 0.25% Triton X-100, and 5 mM 2-mercaptoethanol at pH 7.9), and washed with 1 mL of 0.5x oligo-dT-wash buffer. The mRNA-displayed peptides were eluted with 4.5 mL water plus 5 mM 2-mercaptoethanol into tubes containing Triton X-100 and bovine serum albumin  
10 (BSA, New England Biolabs, Beverly, MA) at final concentrations of 0.15% and 15 µg/mL, respectively.

The mRNA-displayed peptides that eluted from the oligo-dT cellulose column were further purified on Ni-NTA agarose, which binds to the hexahistidine tags on the polypeptides, to remove any mRNA not fused to  
15 polypeptides. The eluted fractions from the oligo-dT cellulose purification were exposed to 0.5 mL Ni-NTA-agarose (Qiagen, Valencia, CA) in Ni-binding buffer [6 M guanidinium chloride, 0.5 M NaCl, 100 mM sodium phosphate, 10 mM Tris(hydroxymethyl)aminomethane, 0.1% Triton X-100, 5 mM 2-mercaptoethanol, 4 µg/mL tRNA (Boehringer-Mannheim, Indianapolis, IN), and  
20 5 µg/mL BSA at pH 8.0)] and incubated for 30 minutes at room temperature. The matrix was then drained, washed with 12 column volumes Ni-binding buffer, and eluted with the same buffer plus 100 mM imidazole. Eluted fractions were combined and de-salted using two successive NAP columns (Amersham-Pharmacia, Piscataway, NJ) equilibrated in 1 mM  
25 Tris(hydroxymethyl)aminomethane, 0.01% Triton X-100, 50 µM EDTA, 0.5 mM 2-mercaptoethanol, 0.5 µg/mL tRNA (Boehringer-Mannheim, Indianapolis, IN), and 50 µg/mL BSA at pH 7.6).

The mRNA portion was then reverse transcribed using Superscript II (Gibco BRL, Rockville, MD) according to the manufacturers instructions, except  
30 that the mRNA concentration was about 5 nM and the enzyme concentration was

1 U/ $\mu$ L. To ensure a high yield in the reaction, a mixture of two primers were used: 1  $\mu$ M of "splint" from the splinted ligation (Cho *et al.*, *supra*), and 1  $\mu$ M of the 3' PCR primer. After 30 minutes at 42°C, the temperature of the reaction mixture was raised to 50° for 2 minutes, and then cooled over 5 minutes to room temperature to allow gradual peptide folding. Finally, the contents were de-salted using NAP columns and subjected to scintillation counting. By comparing the <sup>35</sup>S counts of the purified, reverse transcribed mRNA-peptide fusions to the <sup>35</sup>S-methionine stock and taking into consideration the total methionine concentration in the translation reaction (10  $\mu$ M), the number of displayed peptides in this sample was determined to be  $6.7 \times 10^{12}$ . This number also represents the complexity of the library, since it contained virtually no redundancy (the complexity of puromycin-linker template used in the translation exceeds the number of recovered displayed peptides by a factor of about 35).

Samples from the synthesis and purification of the mRNA displayed peptides were run on an SDS-PAGE gel, as shown in Fig. 1B.

#### Selection of Streptavidin-Binding Peptides

For selection of peptides with high affinity for streptavidin, the above mRNA displayed peptide library was incubated with immobilized streptavidin (Ultralink immobilized streptavidin plus, about 4 mg/mL; Pierce, Rockford, IL) in streptavidin-binding buffer under reducing conditions (40 mM Tris(hydroxymethyl)aminomethane, 300 mM KCl, 2 mM EDTA, 0.1% Triton X-100, 5 mM 2-mercaptoethanol, 100  $\mu$ g/mL BSA, and 1  $\mu$ g/mL tRNA at pH 7.4). The amount of gel used was 0.5 mL in a total volume of 5.5 mL. After incubating for 20 minutes at room temperature, the contents were loaded onto a disposable chromatography column, drained, washed with 14 column volumes of streptavidin-binding buffer, and eluted with five successive aliquots, at 10 minutes intervals, of streptavidin-binding buffer plus 2 mM D-biotin (Sigma, St. Louis, MO) (Fig. 1A). The fraction of the library that survived this purification was 0.08%. Elution fractions were combined, de-salted on NAP columns, and

then PCR-amplified to regenerate the double-stranded DNA library using the described conditions and primers in a 8 mL reaction mixture (Cho *et al.*, *supra*).

This concluded the first round of selection, and the remaining six rounds followed the same protocol except that the translation was scaled down 10-fold, and the number of column volumes for washing the streptavidin column was increased (32 volumes for round 2; 40 volumes for rounds 3, 4 and 6; and 25 volumes for rounds 5 and 7). The streptavidin-binding selection for rounds 5 and 7 was performed directly on the streptavidin-column eluate from the preceding selection rounds, without intervening amplification (the biotin was removed by three successive passages through NAP columns). PCR products amplified after the seventh selection round were cloned using the TOPO TA cloning kit (Invitrogen, Carlsbad, CA), following the manufacture's protocol. The fraction of the library that bound and eluted from the streptavidin column increased in each round, reaching 61% at round seven (Fig. 2A).

#### Characterization of the Selected Library

The eluate from the seventh round of selection was amplified by PCR. The resulting PCR DNA was used to synthesize a library of displayed peptides to confirm that the displayed peptides, rather than the RNA or DNA portion of the library constructs, were responsible for the interaction with streptavidin. Treatment of the library with RNase A did not reduce the extent of binding/elution from the matrix (Fig. 2B). Also, biotin-saturated streptavidin showed no binding to the peptide library (Fig. 2B). These results demonstrated that the interaction of the selected peptides with the streptavidin matrix was specific for the unligated protein, rather than for any other component of the matrix.

### Sequence Analysis of Selected Peptides

Thirty-three randomly chosen clones from the PCR DNA from round seven were chosen for sequencing. Twenty different sequences were observed (Fig. 3). Surprisingly, all 20 sequences were frame-shifted from the intended frame (frame 1) to frame 3 by deletion of two nucleotides or addition of one nucleotide. The designed pattern of polar and non-polar residues was therefore discarded, leaving an unpatterned, essentially random sequence. Prior to the selection, about half of the library members were in frame 1 throughout their entire open reading frames (Cho *et al.*, *supra*). Frame 3 appears to have been enriched over frame 1 due to the increased frequency of the sequence HPQ. Frame 1 has a low incidence (1:45,000 library members) of the sequence HPQ due to the designed polar/non-polar pattern. By contrast, frame 3 had a much higher expected incidence of the HPQ sequence (1:64), similar in frequency to that of a library of the same length and with equal mixtures of all four nucleotides at each position (1:193). Also, frame 3 was rich in histidine, thus allowing retention on the Ni-NTA column. The Ni-NTA purification protocol was intended to eliminate library mRNA molecules not displaying peptide, but was not performed under sufficiently stringent conditions so as to eliminate peptides with small numbers of histidines. Frame 2 had a high incidence of stop codons.

Nineteen of the 20 clones had at least one HPQ motif, and five clones contained two such motifs (Table 1). The clones were organized according to the number of times the HPQ and related tripeptide motifs occur (Table 1). The number of amino acids between the two motifs, when present, ranged from four to 74.

### Binding Affinities of Peptides

To rapidly assay each of the 20 selected peptides to determine their affinity for streptavidin, a new method for preparing, tagging and purifying the peptides was employed. For generation of the DNA-tagged peptides, plasmids containing single inserts were used as templates for PCR-amplification using the same 5'

PCR primer as described for the library construction (Cho *et al.*, *supra*), and a new 3' primer (5'-ATAGCCGGTGCCAAGCTTGCAGCCGCCAGACCAGT-3'; SEQ ID No. 30), which altered the 3' RNA sequence to ACUGGUCUGGCGGCUGCAAGCUUGGCACCGGCUAU (SEQ ID No. 31).

- 5 This sequence was designed to anneal to the photo-crosslinking linker, which has the sequence 5'-psoralen-TAGCCGGTG-A17-CC-puromycin-3', in which the underlined bases are 3'-methoxy nucleotides and the remaining bases are deoxynucleotides (the oligonucleotide was synthesized using reagents from Glen Research, Sterling, VA). This new primer changed the constant C-terminal
- 10 peptide sequence from WSGGCHHHHHHSSA (SEQ ID No. 32) to WSGGCKLGTGY (SEQ ID No. 33), of which the last three amino acids may not be translated because they are annealed to the linker. Each DNA template was transcribed and gel purified as described (Cho *et al.*, *supra*), and then incubated with the psoralen linker under the following conditions: 2  $\mu$ M mRNA, 4  $\mu$ M
- 15 linker, 50 mM Tris(hydroxymethyl)aminomethane, 200 mM KCl, and 10 mM spermidine at pH 7.4 and 70°C for 2 minutes, and then cooled to 4°C over 5 minutes. Samples were then placed in the cold room in a 96 well plate (50  $\mu$ L/well), one inch above which was suspended a UV lamp (366 nm, Ultraviolet Products, Inc., San Gabriel, CA, model number UVL-21) for 15 minutes. Then,
- 20 the reactions mixtures were de-salted using a G-50 Sephadex spin column (Boehringer Mannheim, Indianapolis, IN). The translation/display reactions and oligo-dT-purification were carried out as above. Finally, RNase A (200 ng/mL, 10 minutes, room temperature) was added to degrade the mRNA, leaving peptides fused to a short DNA oligonucleotide. Complete degradation was
- 25 confirmed by SDS-PAGE analysis.

The resulting purified DNA-tagged peptides (DTP) were analyzed in a streptavidin column-binding assay, in which ~500 pM  $^{35}$ S-labeled DTP were mixed with 50  $\mu$ L of the streptavidin matrix in streptavidin-binding buffer, in a total volume of 300  $\mu$ L, and incubated for 10 minutes at room temperature with

30 agitation. Then, the contents were loaded onto a chromatography column. The



column was drained and washed with 80 column volumes of streptavidin-binding buffer, and then eluted with three consecutive aliquots (3 column volumes each) of streptavidin-binding buffer plus 2 mM biotin over a 15 minute period. All fractions (flow-through, washes, elutions, and irreversibly bound counts) were analyzed by scintillation counting to determine the fraction of DTP that bound streptavidin and eluted with biotin (Table 1). The non-selected clone in which two HPQ motifs (separated by 19 amino acids) were introduced encoded the sequence

MDEAHPQAGPVDQADARLVQQGALOHHPQGDRMMSGGCKLGTGY  
(SEQ ID No. 34), in which the underlined portions are identical the HPQ regions of clone SB2.

The results of this analysis are shown in Table 1. For comparison, two HPQ motifs, separated by 19 residues, were introduced into a control, unselected member of the library. The low percentage of this control peptide that specifically eluted from the streptavidin column (0.16%) indicated that the presence of two HPQ motifs was not sufficient for high affinity binding. In contrast, a greater percentage of the selected peptides (8.3 to 88%) was retained on the column during the washing step and then specifically eluted with biotin.

The dissociation constants of the selected peptides for streptavidin were measuring using an electrophoretic mobility shift assay (EMSA). In this assay, DTP's were incubated with varying amounts of pure streptavidin (Pierce Immunopure Streptavidin, Rockford, IL) in streptavidin-binding buffer plus 5% glycerol to increase the density of the solution so that it could collect at the bottom of the gel well. After incubating at room temperature for 20 minutes, the reactions mixtures were moved into the cold room, where they remained for 10 minutes before being carefully loaded onto a 10% polyacrylamide:bisacrylamide (37.5:1, National Diagnostics, Atlanta, GA) gel (thickness 0.7 mm, height 16 cm, width 18 cm) containing 2X TBE, 0.1% Triton X-100 and 5% glycerol. The gel, which had been pre-run for 30 minutes at 13 watts, and the running buffer were pre-cooled to 4°C. Then, the gel was run in the cold room at 13 watts, which

increased the temperature of the gel to about 20°C. The gel was run for 45 to 120 minutes, depending on the mobility of the particular DTP. Then, the gel was fixed in 10% acetic acid and 10% methanol for 15 minutes, transferred to electrophoresis paper (Ahlstrom, Mt. Holly Springs, PA), dried, and analyzed  
5 using a PhosphorImager (Molecular Dynamics, Sunnyvale, CA).

The short DNA oligonucleotide tag on the DTP's allowed them to migrate in a native gel, and the addition of unlabeled ligand (i.e., streptavidin) caused a mobility shift for several of the clones. The concentration of DTPs was less than 1 nM in each titration, and thus the dissociation constant ( $K_d$ ) can be  
10 approximated by the concentration of streptavidin that results in half of the DTP being mobility-shifted. To determine the  $K_d$ , several different measurements were taken in the range of 25-75% of DTP bound (values outside of this range were unreliable due to background and close proximity of the bound and unbound bands in the gel). The  $K_d$  was determined using the equation  $K_d =$   
15  $[\text{streptavidin}] \cdot R$ , where R is the ratio of unbound to bound DTP (ratio of unshifted to shifted band). Independent measurements on gels prepared at different times were used for each clone (the number of different measurements, n, is shown in Table 1). Streptavidin concentrations were measured by UV<sub>282</sub>, using the molar extinction coefficient of 57,000 per monomer.

20 Examples of these mobility shifts in the presence of streptavidin are shown in Fig. 4A. Some clones showed either no shift or poorly defined bands, suggesting that the lifetime of these complexes was too short for detection using this method. We chose five of the most well behaved clones and quantitatively examined their mobility shifts in response to a range of streptavidin  
25 concentrations. An example of a streptavidin titration experiment for peptide SB19 is shown in Fig. 4B, and the data is graphed in Fig. 4C. The dissociation constants for the clones ranged from 110 nM to less than 5 nM (Table 1). These surprisingly high affinities were comparable to those for monoclonal antibody-antigen interactions, demonstrating that even random, non-constrained peptide

libraries can be a source of avid ligands to proteins that do not normally function in peptide binding.

#### Dissection of Clone SB19

- 5           Clone SB19 possessed only one HPQ motif, bound to 85% in the column binding assay, and had a  $K_d$  for streptavidin of 10 nM (Table 1). A series of C-terminal truncation constructs (C1-C4) were constructed and assayed in the streptavidin column-binding assay (Fig. 5). C-terminal truncation analysis of clone SB19 was performed using standard methods by amplifying the clone with  
10 the original 5' primer and a series of 3' primers that truncated the sequence at various positions and also replaced two codons (encoding Asp and Trp) in the C-terminal constant region with methionine codons to increase the  $^{35}\text{S}$ -incorporation. Analogous primers were used for the N-terminal truncation analysis, except that no change was made in the N-terminal constant sequence.
- 15           Deletion of up to 56 residues had no observable effect on the binding strength. Peptide SB19-C4 retained only the first 38 residues from the selected construct (plus the C-terminal MMSGGCKLG sequence, SEQ ID No. 36). Mutating the HPQ motif to HGA reduced the activity by three orders of magnitude (compare construct C4 to M1). Results from the N-terminal  
20 truncation constructs (N1-N3) suggested that binding determinants were spread throughout the N-terminal 38 residues of peptide SB19. Of the peptides tested, SB19-C4 was therefore the minimal peptide retaining full activity in this assay. EMSA analysis of peptide SB19-C4 confirmed high affinity streptavidin-binding, but a fraction (13%) of the peptide was inactive even at streptavidin  
25 concentrations  $>1 \mu\text{M}$ . The majority (87%), however, had an apparent  $K_d$  of 4.9 nM after correction for the amount of inactive peptide.

### Purification of Fusion Protein Containing Streptavidin-Binding Peptide

A fusion protein containing the first 38 amino acids of the SB19-C4 streptavidin-binding peptide (Fig. 7A) was expressed in *E. coli* and then purified from the cell lysate. For the expression of the fusion protein, BL21 (DE3) cells were transformed with a plasmid containing a Maltose Binding Protein--Streptavidin-binding Peptide--His<sub>6</sub>—Protein of Interest insert which encodes a fusion protein containing, from the amino- to carboxy-terminus, maltose-binding protein, the first 38 residues of the SB19-C4 sequence, a hexahistidine tag, and another peptide called 2r18-19dN (Figs. 6A-6C). This insert was constructed using standard molecular biology techniques (see, for example, Ausubel *et al.*, *supra*). Each of these domains of the fusion protein is separated by a few amino acids to allow proper folding of the domains.

A kanamycin-resistant colony was selected and grown overnight in 10 ml LB media with 50 mg/liter kanamycin at 37°C. This starter culture was diluted 100-fold into 1000 ml LB with 50 mg/liter kanamycin, and the culture was grown at 37°C to OD<sub>600</sub> of 1.8 at 37°C. Expression of the fusion protein was induced by addition of 1mM IPTG, and the culture was grown for another two hours. The cells were pelleted by centrifugation at 5000 X g for 20 minutes. The pelleted cells were resuspended in 5% of the original volume of 1 mM EDTA and MBP buffer (10 mM HEPES.HCl, 10 mM HEPES.Na<sup>+</sup>, 200 mM KCl, 0.25% w/w Triton X-100, and 10 mM BME at pH 7.4) and frozen slowly at -20°C overnight. The sample was thawed in the morning and sonicated on ice. The cell lysate was obtained by collection of the supernatant after centrifugation at 14,000 X g for 20 minutes at 4°C.

To purify the fusion protein, the cell lysate was applied to a column containing immobilized streptavidin, with a capacity of about 1 mg/ml, that had been washed with eight column volumes of MBP buffer. Then, the column was washed with 12 column volumes of MBP buffer. The fusion protein was eluted with MBP buffer containing 2 mM biotin. Samples of the cell lysate and eluted protein were analyzed by SDS-PAGE on an 8% gel (Fig. 7B). The lane

containing the purified protein had a band of the expected size. No other bands were observed, except for a faint band of slightly higher mobility (Fig. 7B). This band was probably a degradation product of the fusion protein that was missing a few amino acids from either the amino- or carboxy-terminus but retained the streptavidin-binding peptide and thus retained the ability to bind the streptavidin column. Thirty percent of the fusion protein loaded onto a column containing immobilized streptavidin was recovered after washing the column with 12 column volumes of buffer. Thus, the high affinity of the fusion protein for streptavidin allowed extensive washing of the column to remove contaminating proteins, while retaining a significant amount of the desired fusion protein.

Similar attempts to purify the same protein from the cell lysate, using either amylose resin, which binds to the maltose-binding protein portion of the fusion protein, or Ni-NTA resin, which binds to the hexahistidine tag, resulted in the same expected size band. However, several contaminating bands were also present.

#### Biacore Analysis of the Affinity of a Fusion Protein for Streptavidin

Another fusion protein containing the first 38 amino acids of the SB19-C4 streptavidin-binding peptide was expressed and purified from *E. coli*. This fusion protein contained, from the amino- to carboxy- terminus, maltose-binding protein, the first 38 amino acids of the SB19-C4 sequence, and a hexahistidine tag (Fig. 9B, SEQ ID No. 41). The plasmid (Fig. 9A, SEQ ID No. 40) encoding this fusion protein was constructed using standard molecular biology techniques and used to express the fusion protein in *E. coli*. as described above. This fusion protein was purified from the *E. coli* extract using amylose resin to bind the maltose-binding protein portion of the fusion protein and then Ni-NTA resin to bind the hexahistidine tag.

To measure the affinity of the fusion protein for streptavidin, the fusion protein was immobilized on a biacore chip through the crosslinking of free amino groups in the fusion protein to the biacore chip. Buffer containing streptavidin was washed over the chip, allowing streptavidin to bind the immobilized fusion protein (Fig. 10B). This resulted in an increase in the biacore response units which are proportional to the amount of streptavidin adhering to the biacore chip. Then buffer without streptavidin was washed over the chip, and the biacore response units decreased as streptavidin dissociated from the immobilized fusion protein (Figs. 10A and 10B). To measure the association rate for the binding of streptavidin to the fusion protein, streptavidin concentrations of 1.6, 0.8, 0.4, 0.2, 0.1, or 0.05  $\mu\text{M}$ , (lines "a" to "f" in Fig. 10B, respectively) were washed over the biacore chip. The buffer also contained 40 mM Tris(hydroxymethyl)aminomethane, 300 mM KCl, 2 mM EDTA, 0.1% w/v Triton X-100, and 5 mM 2-mercaptoethanol at pH 7.4. This data was used to calculate an association rate,  $k_a$ , of  $5 \times 10^4/\text{M/s}$ , as described previously (BIACORE X Instrument Handbook, version AA, Biacore AB, Uppsala Sweden, 1997). To measure the dissociation rate, a pulse of 23, 11.5, or 5.75  $\mu\text{M}$  streptavidin in the buffer described above was administered, and then buffer without streptavidin was washed over the chip (Fig. 10A). This data was used to calculate an upper limit of  $2 \times 10^{-3}/\text{s}$  for the dissociation rate,  $k_d$  (BIACORE X Instrument Handbook, *supra*). Based on these calculated association and dissociation rates, the dissociation constant,  $K_d$ , for the binding of streptavidin by this fusion protein was less than 40 nM. This result confirms the high affinity binding of the SB19-C4 peptide for streptavidin that was observed in the streptavidin column-binding assay and the EMSA assay (Table 1). Additionally, this result demonstrates that this peptide maintains its high affinity for streptavidin when expressed as part of a fusion protein.

Other Embodiments

From the foregoing description, it will be apparent that variations and modifications may be made to the invention described herein to adopt it to various usages and conditions. Such embodiments are also within the scope of  
5 the following claims.

All publications mentioned in this specification are herein incorporated by reference to the same extent as if each independent publication or patent application was specifically and individually indicated to be incorporated by reference.

10

What is claimed is:

Claims

1. A peptide which binds streptavidin with a dissociation constant less than  $10\ \mu\text{M}$ , wherein said peptide is not disulfide bonded or cyclized.
- 5        2. A peptide which binds streptavidin with a dissociation constant less than  $10\ \mu\text{M}$ , wherein the amino acid sequence of said peptide does not contain an HPQ, HPM, HPN, or HQP motif.
- 10       3. A peptide which binds streptavidin with a dissociation constant less than  $23\ \text{nM}$ .
4. The peptide of claim 1 or 2, wherein said dissociation constant is less than  $1\ \mu\text{M}$ .
- 15       5. The peptide of claim 4, wherein said dissociation constant is less than  $100\ \text{nM}$ .
6. The peptide of claim 5, wherein said dissociation constant is less than  $50\ \text{nM}$ .
- 20       7. The peptide of claim 3, wherein said dissociation constant is less than  $10\ \text{nM}$ .
8. The peptide of claim 7, wherein said dissociation constant is less than  $5\ \text{nM}$ .
- 25       9. The peptide of any one of claims 1-3, comprising at least 10 consecutive amino acids of any one of SEQ ID Nos. 1-29.



10. The peptide of claim 9, comprising at least 25 consecutive amino acids of one of SEQ ID Nos. 1-29.

11. The peptide of claim 10, comprising at least 50 consecutive amino acids of any one of SEQ ID Nos. 1-29.

12. The peptide of claim 11, comprising at least 100 consecutive amino acids of any one of SEQ ID Nos. 1-29.

13. The peptide of any one of claims 1-3, comprising the amino acid sequence of any one of SEQ ID Nos. 1-29 or 35.

14. A nucleic acid encoding a peptide of any one of claims 1-3.

15. A vector comprising a nucleic acid of claim 14.

16. A fusion protein comprising a protein of interest covalently linked to:

(a) a peptide which binds streptavidin with a dissociation constant less than 10  $\mu$ M, wherein said peptide is not disulfide bonded or cyclized;

(b) a peptide which binds streptavidin with a dissociation constant less than 10  $\mu$ M, wherein said peptide does not contain an HPQ, HPM, HPN, or HQP motif; or

(c) a peptide which binds streptavidin with a dissociation constant less than 23 nM.

17. The fusion protein of claim 16, wherein said peptide is attached to the amino terminus or the carboxy terminus of said protein of interest, or wherein said peptide is positioned between the amino and carboxy termini of said protein of interest.

18. The fusion protein of claim 16, wherein said peptide is linked to said protein of interest by a linker comprising a protease-sensitive site.

19. A nucleic acid encoding a fusion protein of claim 16.

5

20. A vector comprising a nucleic acid of claim 19.

21. A method of producing a streptavidin-binding fusion protein, said method comprising the steps of:

10 (a) expressing in a host cell, a gene encoding a fusion protein of claim 16; and

(b) culturing said host cell under conditions appropriate for production of said fusion protein.

15 22. A method of purifying a protein of interest from a sample, said method comprising the steps of:

(a) expressing in said sample, a fusion protein comprising said protein of interest covalently linked to:

20 (i) a peptide which binds streptavidin with a dissociation constant less than  $10 \mu\text{M}$ , wherein said peptide is not disulfide bonded or cyclized;

(ii) a peptide which binds streptavidin with a dissociation constant less than  $10 \mu\text{M}$ , wherein said peptide does not contain an HPQ, HPM, HPN, or HQP motif; or

25 (iii) a peptide which binds streptavidin with a dissociation constant less than  $23 \text{ nM}$ ;

(b) contacting said sample with streptavidin under conditions that allow complex formation between said fusion protein and said streptavidin;

(c) isolating said complex; and

30 (d) recovering said fusion protein, thereby purifying said protein of interest from said sample.

23. A method of detecting the presence of a fusion protein of claim 16 in a sample, said method comprising the steps of:

(a) contacting said sample with streptavidin under conditions that allow complex formation between said fusion protein and said streptavidin;

5 (b) isolating said complex; and

(c) detecting the presence of said streptavidin, wherein the presence of said streptavidin indicates the presence of said fusion protein in said sample.

24. The method of claim 23, wherein step (c) comprises detecting  
10 the presence of said streptavidin in said complex

25. The method of claim 23, wherein step (c) comprises detecting the presence of said streptavidin recovered from said complex.

15 26. The method of claim 23, wherein step (c) further comprises measuring the amount of said streptavidin, wherein the amount of said streptavidin is correlated with the amount of said fusion protein in said sample.

1/18

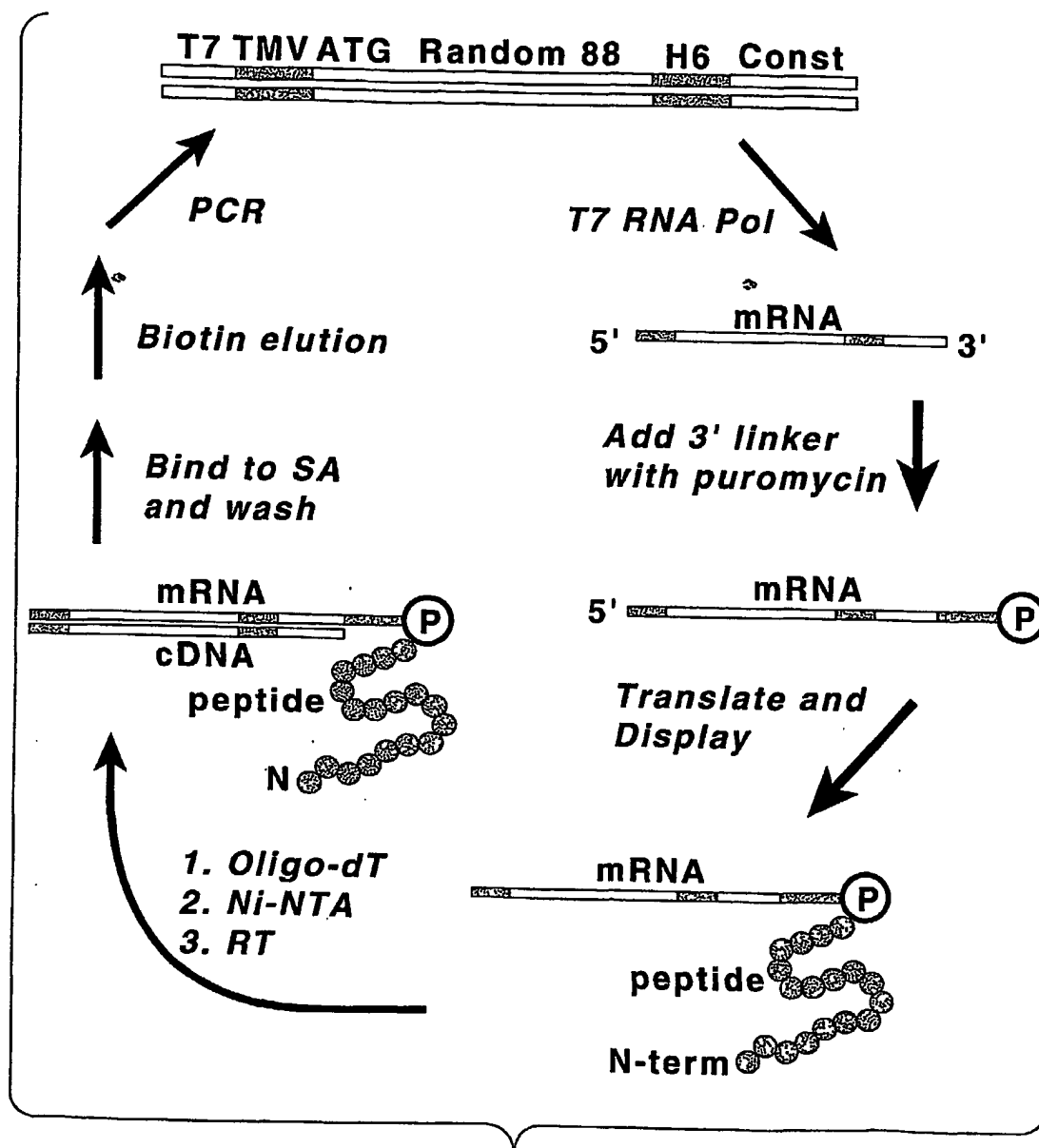


Fig. 1A

SUBSTITUTE SHEET (RULE 26)

2/18

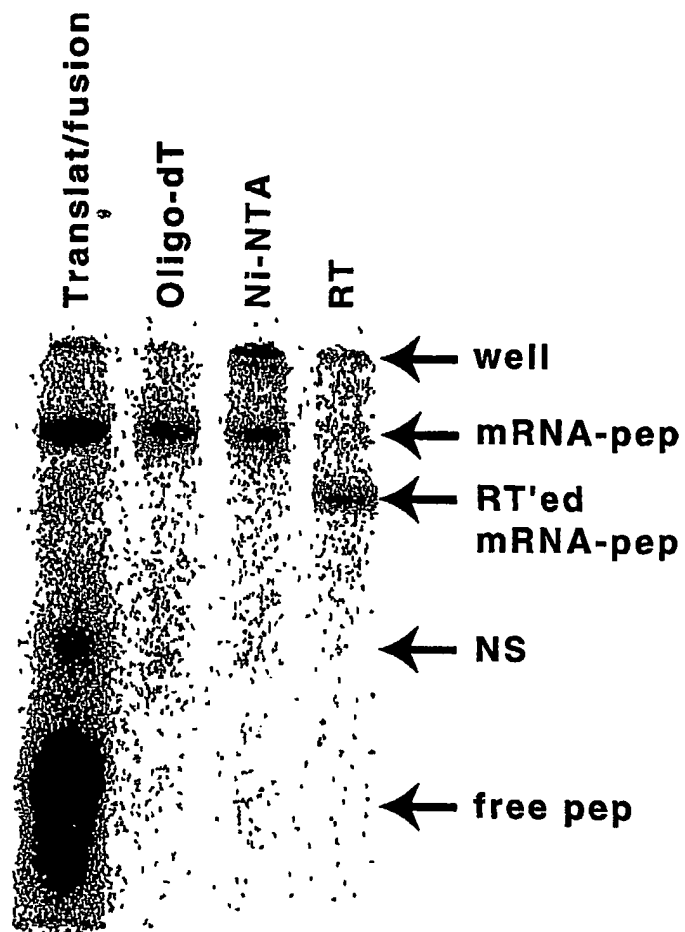


Fig. 1B

SUBSTITUTE SHEET (RULE 26)

3/18

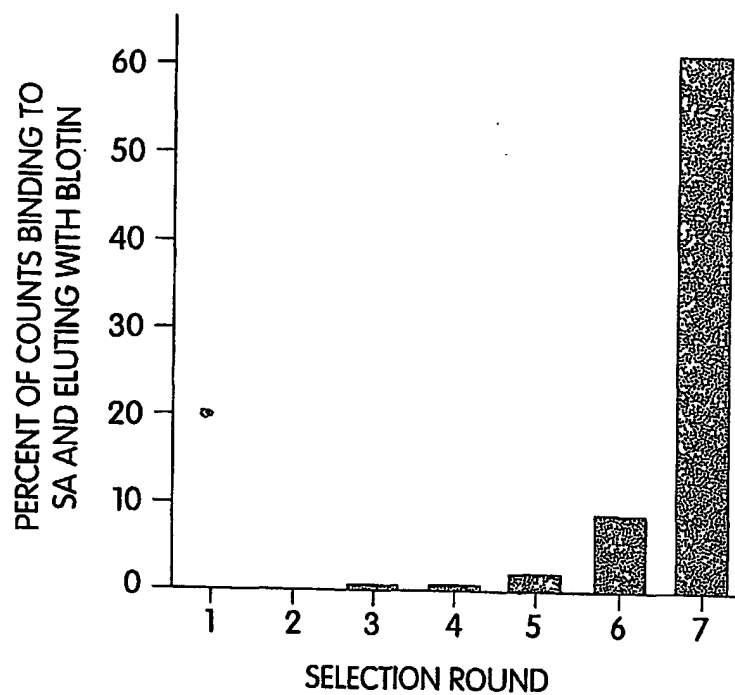


Fig. 2A

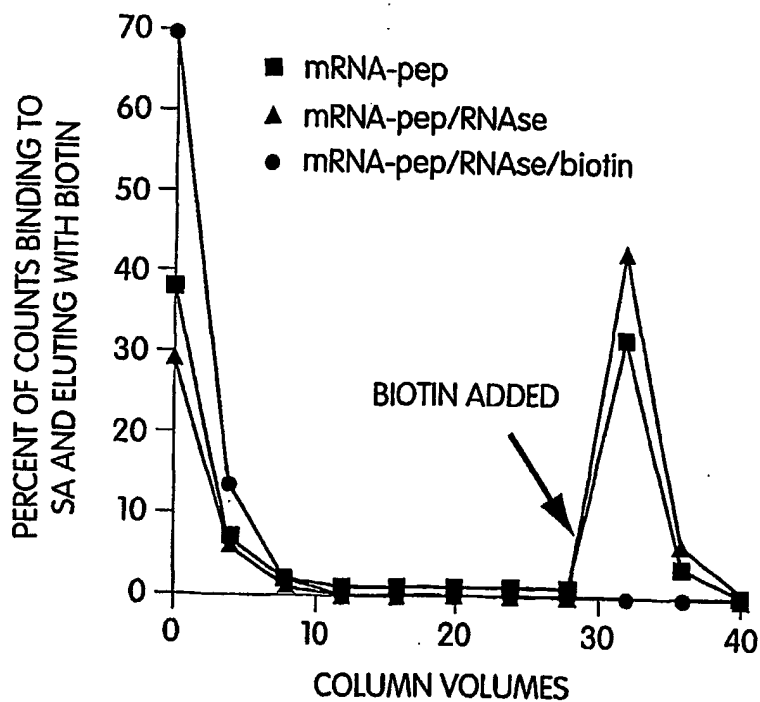


Fig. 2B

SUBSTITUTE SHEET (RULE 26)

name #	SEQ ID NO:
SB1 3	<u>MDEKTHCTISMGAVPLVPHHPQGDPLRLRRPQALLVREHPQGDVALVEHHEGVDRGLVALPELHAELGEPVGDVQGPVEQVQGVVDALVWRLPPS</u> 1
SB2 2	<u>MDEKTHCFHPGDHLVRLVEELQALAEGLRQGRQPHRLPRRRPHHLQLLLDEAHPQAGPLRERAHQVDGRLLQLHHPQCDRLLOQPQDHPLELVMRLPPS</u> 2
SB3 4	<u>MTRRPTASSSSCVRHLLLRQGEHGHQALEDRDKARHVRVVEGVVEVIGGLDRLARARHEALHPQAGLVHLPLHGGDLGGHRLRLVLEAHPQCDRLGLAVHHH</u> 3
SB4 1	<u>MDEKTHWGLSTWRGEPPLHHPQAGRLPLDRRRARHRRILGAEPGGVDHGLRLELLDDHRPLVPDHPQGRGPLQRGDLPQVVPVLRRLRHAHVLCGLGLAAATIT</u> 4
SB5 3	<u>MDEKTHWVNVVHPQGDLLVRGCHDVEALHDQGLHQLDLLVGPPPEVWRALRGVIGGLRRLVPLDHPQGEALDQARQRPQHLLLELHHRALPPALVWRLPPS</u> 5
SB6 1	<u>MDEKTHWLNANFEELARLDGLREGEDHPLVLRHHPQGDGLLDQPLGRHRLDGEVREGDRPLDQCGEEDLGALVDDGEVLDGLVHVGVVHVDPLVCGCHHH</u> 6
SB7 1	<u>MDEKTHWFGTLNSFPPTHWMSAVGNKIDCSFNWNLNLNHWLSSCHPDGALDDQHPQGDALVGRDDGVVQALRLEGQHQRRLAQRRAADRHRQLVWRLPPS</u> 7
SB8 1	<u>MDEKTHCTIETLNFSTHWWKLHHPQGDALLDDGVRPHHPLADEGGGLDQGLGHRRGVWAERLARRDPEVLEGLVERHGLVPRLRHGGERHAEPVWRLPPS</u> 8
SB9 1	<u>MDEKTHCNTGLYDGAADCFNELNKDVAPLVEGRHDLVEGLLLEHHPQGDPLVAHRQLVHHPLGRGERHRRALVPQGEHQPRLQPVVDLGRRLVWRLPPS</u> 9
SB10 1	<u>MDEKTHWHERAQELVGGLLLDHHPQRLLEPRGPRPLRGLVHERGHQPPPLAGRVEEADGGLLRDGGGELEPLVREGEDHLEPLDDELDAQPRGLVWRLPHHH</u> 10
SB11 1	<u>MDEKTHWHERVHHLADGLEQHPQGORRPLVERHRQVPRGLVRELQHEGLPLEHPAGVHVIRLHQGDDRDVDGLVDGHRDVRGLEREVEGDPHRLVWRLPPS</u> 11
SB12 4	<u>MDKDLLLEELERELVHHPQGGLLPLRGQVGHDAERLGAEVDDLRGGLLDEPQRAVAGLHHVPHRVGQRLVHEVRELDGLLDQRDDLRQRLVWRLPPS</u> 12
SB13 2	<u>MEREDPLDEQLRELREALVDEHPQGGQAQALHRHDGGEHVPLRRVQHRLQPGIQLHLEPQPLGLLGLQARLQPLAGEHEGDGAGLQRPVGHQGRRLVWRLPPS</u> 13
SB14 1	<u>MDEKTHRTLSVLSFNDWLQTKACWRLVEGLHHPQGLVREHEVDVLPLAEVQVVGGLADGVEQPGGGLLHRAQRVDHPLPDHAGQVLRGLVWRLPPS</u> 14
SB15 1	<u>MDEKTHWLEDLKGVLKDCIKOLMDFTKDCRSRVPQPLLHHDGRGEPVPLLEAGRDGLGLPRAPRQARPLHGHGRHDLHEPLVLQDHPQGGPLVCGCHHH</u> 15
SB16 1	<u>MDEKTHWVLQLHPQGDRLGPRHGGDDVRLVVGQEGVLEGLDGRPRRRRRLPREDEHVRVALVDQVRDLAERLVEEVDDGVEALRHILGLPQDEPRSGGCHHH</u> 16
SB17 2	<u>MDEKTHWVGLQEPGLPHGGVGEVPGGLVLRHHPQDRDLVDGVCPHGRALARRPHRVVEGLHLLLRGGERLPPDGPRLGLLGGELDRADPALVWRLPPS</u> 17
SB18 1	<u>MDEKTHCAVNVNGLTHWCHRVAHLQPLDHPQGDHLRLLEPLGHALVDPLVQGVVEVVRPLQLDVGVRVALVEQVAEVEGGLDHEAGQAAGALVWRLPPS</u> 18
SB19 1	<u>MDEKTHWGRGGHVVEGLAGELEQLRARLEHHPQGGREPLVQVEVDVEGLVQDLHGVAAGLLDPVEKLLTDMFKFKFNKVSCKDKMTFFYLEMYDMSGGCHHH</u> 19
SB20 1	<u>MNEKTHCKLNFKNVNIADWLAEFFHGGGGLGRRDGVVQRLVDGVQERVERLDRDPGLGLRLLELHHRDRRLRGGEHLLRDHPLEPDDHLVVGGLVWRLPPS</u> 20

Fig. 3

5/18

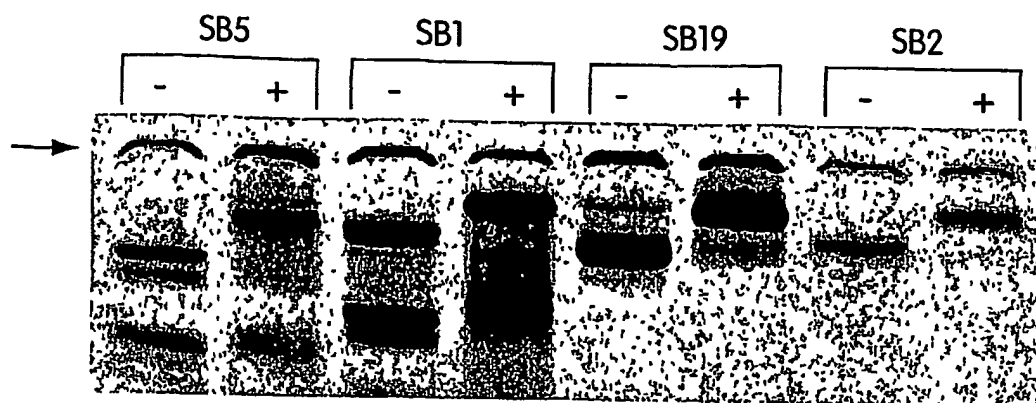


Fig. 4A

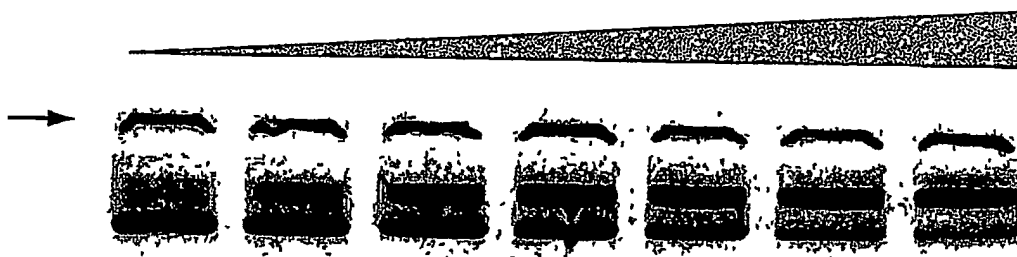


Fig. 4B

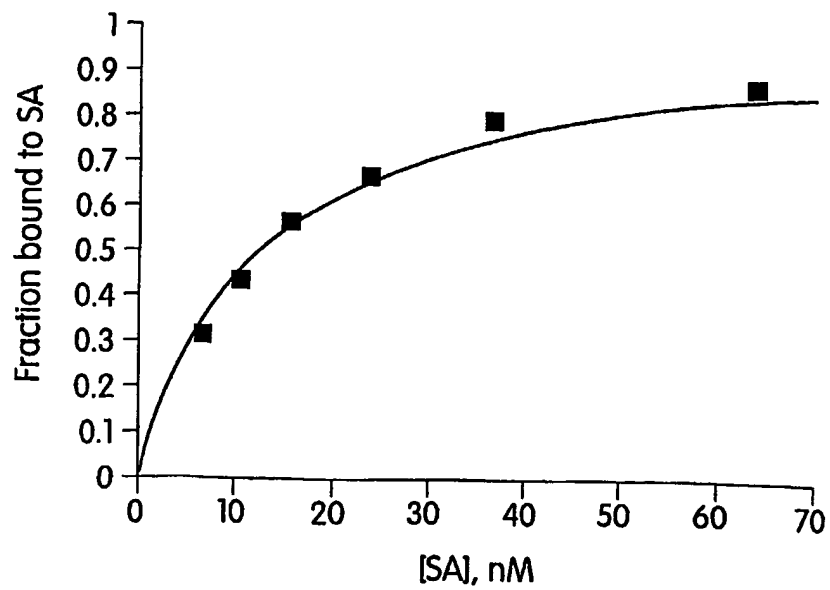


Fig. 4C

SUBSTITUTE SHEET (RULE 26)



		% binding	SEQ ID NO:
FL	MDEKTTGWRGGHVVEGLAGELEQLRARLEHHHPQGQREPLVQEVEDVDEGLVQDLHGVAAGLLDPVEKLLTDWFKFKNVSKDCKMTFYLEMYDWSGGCKLG	85	21
C1	MDEKTTGWRGGHVVEGLAGELEQLRARLEHHHPQGQREPLVQEVEDVDEGLVQDLHGVAAGLLDPVEKLLTDWFKFKNVSKDCKMTFYLEMYDWSGGCKLG	87	22
C2	MDEKTTGWRGGHVVEGLAGELEQLRARLEHHHPQGQREPLVQEVEDVDEGLVQDLHGVAAGLLDPVEKLLTDWFKFKNVSKDCKMTFYLEMYDWSGGCKLG	88	23
C3	MDEKTTGWRGGHVVEGLAGELEQLRARLEHHHPQGQREPLVQEVEDVDEGLVQDLHGVAAGLLDPVEKLLTDWFKFKNVSKDCKMTFYLEMYDWSGGCKLG	89	24
C4	MDEKTTGWRGGHVVEGLAGELEQLRARLEHHHPQGQREPLVQEVEDVDEGLVQDLHGVAAGLLDPVEKLLTDWFKFKNVSKDCKMTFYLEMYDWSGGCKLG	88	25
M1	MDEKTTGWRGGHVVEGLAGELEQLRARLEHHHPQGQREPLVQEVEDVDEGLVQDLHGVAAGLLDPVEKLLTDWFKFKNVSKDCKMTFYLEMYDWSGGCKLG	0.065	26
N1	MD		27
N2	MD		28
N3	M		29

6/18

Fig. 5



8/18

CGCAAAATTGTCGGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTGATGGTAGAAACGAAGCGGGC  
GTCGAAGCCCTGTAAGCGGGGTGCACAAATCTTCTCGCGCAACCGGTCACTGGGTGATCATTAACCTATCCGCTGGATGA  
CCAGGATGCCATTGCTGTGAAGCTGCCAGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACACAGACACCCATCA  
ACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGCGGTGGAGCATCTGGTCGCATTGGGTCAACAGCAAATCGCG  
CTGTTAGCGGGCCCATTAAGTTCTGCTCGCGCGCTCGCTCGGTGCTGGTGGTGGTGGCATAAATATCTCACTCGCAATCAAAT  
TCAGCCGATAGCGGAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTCAACAACCATGCAAAATGCTGAATGAGGGCA  
TCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGTGGCGCAATGCGGCCATTACCGAGTCCGGGCTGCGC  
GTTGGTGGGATATCTCGGTAGTGGGATACGAGATACCGAACGCTGCTGCAACTCTCTCAGGGCCAGCGGTGAAGGGCAATC  
ACAGGATTTTCGCCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGAAGGGCAATC  
AGCTGTTGCCCGTCTCACTGGTGAAAGAAAACCAACCTGCGGCCCAATACGCAAAACCGCTCTCCCGCGCGTTGGCC  
GATTCATTAATGCAGCTGGCACGACAGGTTCCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAAGTTAGC  
TCACTCATTAGGCACCGGATCTCGACCGATGCCCCTGAGAGCCTTCAACCCAGTCACTCCTTCCGGTGGCGCGCGGGC  
ATGACTATCGTCGCCGCACTTATGACTGTCTTCTTATCATGCAACTCGTAGGACAGGTCCGGCATCGCTCTGGTCAAT  
TTTCGGCGAGGACCGCTTTCGCTGGAGCGGACGATGATCGGCCCTGTGCTTGGGTATTCGGAATCTTGACGCGCCCTCG  
CTCAAGCCTTCGTCACTGGTCCGCCCAACCAACGTTTCGGCGAGAAACGAGGCCATTATCGCCGGCATGGCGGCCCCACGG  
GTGCGCATGATCGTGTCTGTTAGGACCCGGCTAGGCTGGCGGGTTCCTTACTGTTAGCAATGAATGAATCAACC  
GATACGCGAGCAACGTGAAGCGACTGCTGCTGCAAAACGCTGCGAACCAGCAACACATGAATGGTCTTCGGTTTCC  
GTGTTTCGTAAGTCTGAAAACGGGAACGAGCCCTGCAACCATATGTTCCGGATCTGCATCGCAGGATGCTGCTGG  
CTACCCGTGGAACACCTACATCTGTATTACGAAGCGCTGGCATTGACCCGTAGTGATTTTTCTTGTTCCCGCCGAT  
CCATACCGCCAGTTGTTACCCTCACAAAGTTCCAGTAACCGGGCATGTTTCATCATCAGTAACCCGTATCGTGAGCATCC  
TCTCTCGTTTCATCGGTATCATTAACCCCATGAACAGAAATCCCTTACACGGAGGCATCAGTGACCAACACAGGAAAA  
ACCGCCCTTAACATGGCCCGCTTATCAGAACCCAGACATTAACGCTTCTGGAGAACTCAACGAGCTGGACGCGGATGA  
ACAGGCAGACATCTGTGAATCGCTTACGACCAACGCTGATGAGCTTACCGCAGCTGCCCTCGCGCTTTCGGTGATGACG  
GTGAAAACCTCTGACACATGCAGTCCCGGACGGTCACAGTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGT  
CAGGGCGGTGTCAGCGGTGTTGGCGGTGTCGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGATATCTGG  
CTTAACATGCGGCATCAGACGATGTACTGAGAGTGCAACCATATATGCGGTGTGAAATACCGCACAGATGCGTAAGG  
AGAAAATACCGCATCAGGCGCTCTTCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCGGCTGCGCGAGCGGT  
ATCAGTCACTCAAAGGCGTAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAACATGTGAGCAAAAAGGCC  
AGCAAAAGGCCAGGAACCGTA AAAAGGCCGCTTGTGCGGTTTTTCCATAGGCTCCGCCCTCCCTGACGAGCATCAAAA  
AATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAAGCGTTTCCCGCTGGAAGCTCCCTCGT  
GCGCTCTCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAACGCTGGCGCTTCTCTCATATA

**Fig. 6A (continued)**

9/18

GCTCAGCTGTAGGTATCTCAGTTCCGTTGAGTTCGTTCCGCTCAAAGCTGGGCTGTGTGCACGAACCCCCCGTTAGCCCC  
GACCGTGGCCCTTATCCGGTAACATATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCAC  
TGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTACGGCTACACTA  
GAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTCGGAAAAGAGTTGGTAGCTCTTGATCCGGCAAA  
CAACCCAGCTGGTAGCGGTGGTTTCTTGGTTGCAAGCAGAGATTACGGCAGAAAAGGATCTCAAGAAAGATCC  
TTTGATCTTTCTACGGGCTCTGACGCTCAGTGAACGAAACTCACGTTAAGGATTTTGGTTCATGAACAATAAACTG  
TCTGTTACATAAACAGTAATAACAAGGGTGTATGAGCCATATTCAACGGGAACGCTTGTCTAGGCCCGATTAA  
TTCCAACATGGATGCTGATTATATGGGTATAAATGGGCTCGGATAATGTGGGCAATCAGGTGCGACAATCTATCGAT  
TGATGGGAAGCCCGATGCCCGAGAGTTGTTCTGAAACATGGCAAGGTAGCGTTGCCAATGATGTTACAGATGAGATG  
GTCAGACTAACTGGCTGACGGAATTTATGCCCTCTCCGACCATCAAGCATTTTATCCGTACTCTGATGATGATGATGGTT  
ACTCAACCTGCGATCCCGGGAACACAGCATCCAGGTATTAGAGAAATATCCTGATTCAGGTGAAAATATTGTTGATG  
CGCTGGCAGTGTCCCTGCGCGGTGCAATTCGATTCCTGTTGTAATTGTCCTTTTAAACAGCGATCGCGTATTTTCGTCTC  
GCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTATTTGATGACGAGCGTAATGGTGGCTGTTGA  
ACAACTCTGGAAAGAAATGCATAAACTTTGCCATTTCCAGGATTCAGTCTCAGTCAATCGAGATTTCTCACTTGATA  
ACCTTATTTTGACGAGGGGAAATTAATAGTTGTAATGATGTTGGACGAGTCGGAATCGCAGACCGATACCGGATCTT  
GCCATCCTATGGAACCTGCCCTGGTGAGTTTCTCCTTCAACAGAACGGCTTTTCAAAAATATGGTATGATAATCC  
TGATATGAATAAATGCAGTTTCAATTTGATGCTCGATGAGTTTCTAAGAAATTAATTCATGAGCGGATACATATTGAA  
TGATTTAGAAAAATAACAAATAGGGGTTCCCGGCACATTTCCCGAAAGTGCCACCTGAAATTTGTAACGTTAATAT  
TTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAATTTTAAACCAATAGGCCGAAATCGGCAAAATCCCTTATA  
AATCAAAAGAAATAGACCGAGATAGGTTGAGTGTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCC  
AACGTCAAAAGGGGAAACCCGTCTATCAGGGCGATGGCCCCACTACGTGAACCATCACCTTAATCAAGTTTGTGGGTC  
GAGTGCCGTAAAGCACTAAATCGAACCCCTAAGGAGCCCCCGATTAGAGCTTGACGGGAAAGCCGGCGAACGTGG  
CGAGAAAGGAAGGAAGCAAGGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACCGTCAAGTGGCGTAACCCACC  
ACACCGCGCGCTTAATGCGCGCTACAGGGCGCTCCCATTCGCCA

Fig. 6A (continued)

10/18

(SEQ ID No.: 38)

MGIEGKLV I W I N G D K G Y N G L A E V G K K F E K D T G I K V T V E H P D K L E E K F P Q V A A T G D G P D I I F W A H D R F G G Y A Q S G L L A E I  
 T P D K A F Q D K L Y P F T W D A V R Y N G K L I A Y P I A V E A L S L I Y N K D L L E N P P K T W E E I P A L D K E L K A K G S A L M F N L Q E P Y F T W P  
 L I A A D G G Y A F K Y E N G K Y D I K D V G V D N A G A K A G L T F L V D L I K N K H M N A D T D Y S I A E A F N K G E T A M T I N G P W A W S N I D T S K  
 V N Y G V T V L P T F K G Q P S K P F V G L S A G I N A A S P N K E L A K E F L E N Y L L T D E G L E A V N K D K P L G A V A L K S Y E E E L A K D P R I A A  
 T M E N A O K G E I M P N I P Q M S A F W Y A V R T A V I N A A S G R Q T V D E A L K D A Q T N S S G G S G S G M D E K T T G W R G G H V V E G L A G E L E Q  
 L R A R L E H P Q Q R E P G S G H H H H H E F L V P R G S M D P C V K C K V A P R N W K V K N K H L R I Y N M C K T C F N N S I D I G D D T Y H G H D D

Fig. 6B

(SEQ ID No.: 39)

M D P C V K C K V A P R N W K V K N K H L R I Y N M C K T C F N N S I D I G D D T Y H G H D D

Fig. 6C

11/18

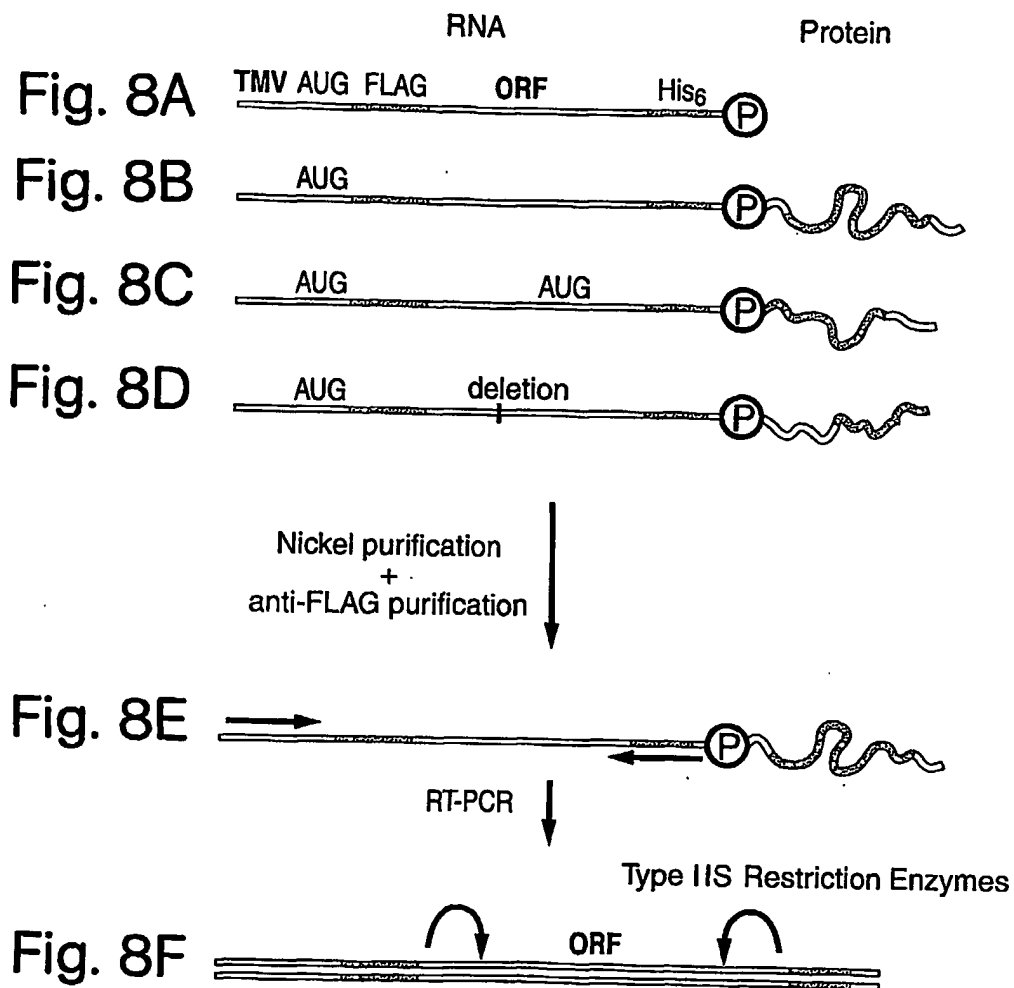
MDEKTTGWRGGHVV
   
 EGLAGELEQLRARLEHHPQQQREP

Fig. 7A



Fig. 7B

12/18

**SUBSTITUTE SHEET (RULE 26)**

[illegible]

**Fig. 9A**



ACTATCCGCTGGATGACACAGGATGCCATTGCTGTGGAAAGCTGCCTGCACATAATGTTCGGCGGTTATTTCTTGATGTCTCT  
 GACCAGACACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATCTGGTCGCATTTGGG  
 TCACCAGCAATCGCGCTGTAGCGGCCCATTAAGTTCTGTCTCGGCGCTGTGCTGTGGCTGGCTGGCATAAATATC  
 TCACCTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTCAACAAACCATGCAA  
 ATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTGGTCCAAACGATACGATGGCGCTGGCGCAATGCGCCCATTAC  
 CGAGTCGGGCTGGCGCTGGTGGCGGATATCTCGGTAGTGGGATACGAGATACCGAAGACAGCTCATGTTATATCCCGC  
 CGTTAACCAACCATCAACAGGATTTTCGCCCTGTGGGCAACACGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAG  
 GCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACACCCCTGGCGCCCAATACGCAACCCGCTC  
 TCCCGCGCGTTGGCCGATTCAATTAATGCAGTGGCACGACAGGTTTCCGACTGGAAGCGGGCAGTGAGCGCAACGCA  
 ATTAATGTAAAGTTAGTCACTCACTATTAGCACCGGGATCTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTT  
 CCGTGGCGCGGGCATGACTATCGTGGCGGACTTATGACTGTCTTCTTATCATGCAACTCGTAGGACAGGTGCCGG  
 CAGCGCTCTGGTCAATTTTCGGCGAGGACCGCTTTCGCTGGAGCGGACGATGATCGGCTGTGCTGTCCGTTATTCGGA  
 ATCTTGACAGCCCTCGCTCAAGCTTCGTCACCTGGTCCGCCACCAACGTTTCGGCGAGAACGAGGCCATTTATCGCCGG  
 CATGGCGGCCCCACGGGTGGCATGATCGTGTCTCGTGGAGCCCGCTAGGTAGGCTGGCGGCTGCTTACTGGTT  
 AGCAATGAATCACCGATACCGGAGCGAACGTAAGCGACTGTCTGCAAAACGTCAGCGCTGCACCATTAATGTTCCGGATCGCAT  
 ATGGCTTTCGGTTTCCGTTTTCGTAAGTCTGGAACCGGAAAGTCAGCGCTGGCATTAACGAAGCGTGGCATTAACCTGATTTT  
 CGAGGATGCTGTGGTACCTGTGGAAACACCTACATCTGTAATTAACGAAGCGTGGCATTAACCTGATGATTTTC  
 TCTGGTCCCGCGCATCCATACCGCCAGTTGTTACCTCACAAAGTTCCAGTAACCGGCGCATGTTTCATCATCAGTAACC  
 CGTATCGTGAGCATCCTCTCTCGTTTCATCGGTATCATACCCCATGAACAGAAATCCCTTACACGGAGGCATCAGT  
 GACCAACAGGAAACACCGCCCTTAACATGGCCCGCTTATCAGAACCCAGACATTAACGCTTCTGGAGAAACTCAACG  
 AGCTGGACGCGGATGAACAGGCAGACATCTGTGAATCGCTTACGACACGCTGATGAGCTTTACCGCAGCTGCCCTCGCG  
 CGTTTCGGTGTAGCGGTGAACCTCTGACACATGACGCTCCCGAGACGGTACAGCTTGTCTGTAAGCGGATGCCGG  
 GAGCAGACAAGCCGTCAGGGCGGCTCAGCGGCTGTGGCGGCTGTGGCGGCGAGCCATGACCCAGTCACGTACGCGATA  
 GCGGAGTGTACTGCTTAACTATGCGGCATCAGAGCAGATTGTAAGAGTGACCATATATGCGGTGTGAATACC  
 GCACAGATGCGTAAGGAGAAATACCGCATCAGGCGCTTCCGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTC  
 GGCTGGCGGAGCGGTATCAGCTCACTCAAGGCGGTAATAAGGTTATCCACAGAAATCAGGGGATAACGCGAGAAAGAAC  
 ATGTGAGCAAAAGGCCAGCAAAAGGCCAGCAACCGTAAAGGCCGCTGTGGCGTTTTCATAGGCTCCGCCCGCC  
 TGACGAGCATCACAAATCGACGCTCAAGTCAGAGGTGGCGAAACCGCAGACACTATAAGATACAGGCGGTTTCCCC  
 CTGGAAGCTCCCTCGTCTCTGTTCCGACCTTACCGGATACCTGTCCGCTTCTCCCTTCCGCTTCCGGAAGC  
 GTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCGCTCAAGCTGGGCTGTGTGACGA  
 ACCCCCCGTTACGCCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTAGTCCAAACCCGGTAAGACACGACTTATCGC

Fig. 9A (continued)

CACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCT  
AACTACGGCTACACTAGAAAGGACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAATAAGAGTTGGTAG  
CTCTTGATCCGGCAACAAACCCGCTGGTAGCGGTGGTTTTTGTGTTGCAAGCAGAGATTACGCGCAGAAAAAAG  
GATCTCAAGAAGATCCTTTGATCTTTCTACGGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTC  
ATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAAGGGTGTATAGCCATATTCACGGGAACGCTCTTGCT  
CTAGGCCGGGATTAAATCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGGATAATGTCGGGCAATCAGGT  
GCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTCTGAAACATGGCAAGGTAGCGTTGCCAATGA  
TGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCCTCTCCGACCATCAAGCATTTTATCCGTACTC  
CTGATGATGCTGTTACTCACCCTGCGATCCCGGGAAAAACAGCATTCAGGTATTAGAAGAATATCCTGATTCAGGT  
GAAAAATATTGTTGATGCGCTGGCAGTGTCTCTGCGCCGGTTGCAATTCGATTCCTGTTTGTAAATGTCTTTTAAACAGCGA  
TCGCGTATTTCTGCTCAGCGCCCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGTATGACGAGCGTA  
ATGGCTGGCCTGTTGAACAAGTCTGGAAGAAATGCATAAACTTTTGCCATTCTCACCGGATTCTAGTCGTCACATCATGGT  
GATTTCTCACTGATAACCTTATTTTGTGACGAGGGGAAATTAATAGGTGTATTGATGTTGGACGAGTCGGAATCGCAGA  
CCGATACAGGATCTTGCCATCCTATGGAACTGCCCTGGTGAGTTTCTCCTTCAATTACAGAAACGGCTTTTCAAAAAT  
ATGGTATTGATAATCCTGATATGAATAAATGCAATTTGATGCTCGATGAGTTTCTAAGAAATTAATTCATGAG  
CGGATACATATTGAAATGATTTAGAAAAATAAACAATAGGGTTCCGCGCACATTTCCCGGAAAGTGCCACCTGAAA  
TTGTAAACGTTAATATTGTTAAATTCGCGTTAAATTTTGTAAATCAGCTCATTTTAAACCAATAGGCCGAAATC  
GGCAAAATCCCTTATAATCAAAAGAAATAGACCGAGATAGGTGAGTTGTTCCAGTTTGGAAACAAGATCCACTATT  
AAAGAACGTGGACTCGAACGTCAAAGGGCGAAACCCGTCTATCAGGGCGATGGCCACTACGTGAACCATCACCTAAT  
CAAGTTTTTGGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCTTAAGGGAGCCCGGATTTAGAGCTTGACGGGA  
AAGCCGGGAACGTGGCGAGAAAGGAAGGGAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCTAC  
GCTGCGCGTAACCAACACACCCCGCGCTTAATGCGCCGCTACAGGGCGGTCCCATTCGCCA

Fig. 9A (continued)

16/18

(SEQ ID No.:41)

MGIEEGKLVIWINGDKYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEI  
TPDKAFQDKLYPFTWDVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIIPALDKELKAKGKSALMFNLQEPYFTWP  
LIAADGGYAFKYENGKYDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSK  
VNYGVTVLPTFEKGQPSKPFVGVLSAGINAASPKNELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAA  
TMENAQKGEIMPNI PQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSGSGSGMDEKTTGWRGHHVVEGLAGELEQ  
LRARLEHHHPQGQREPGSGHHHHHHHEF

Fig. 9B

SUBSTITUTE SHEET (RULE 26)

17/18

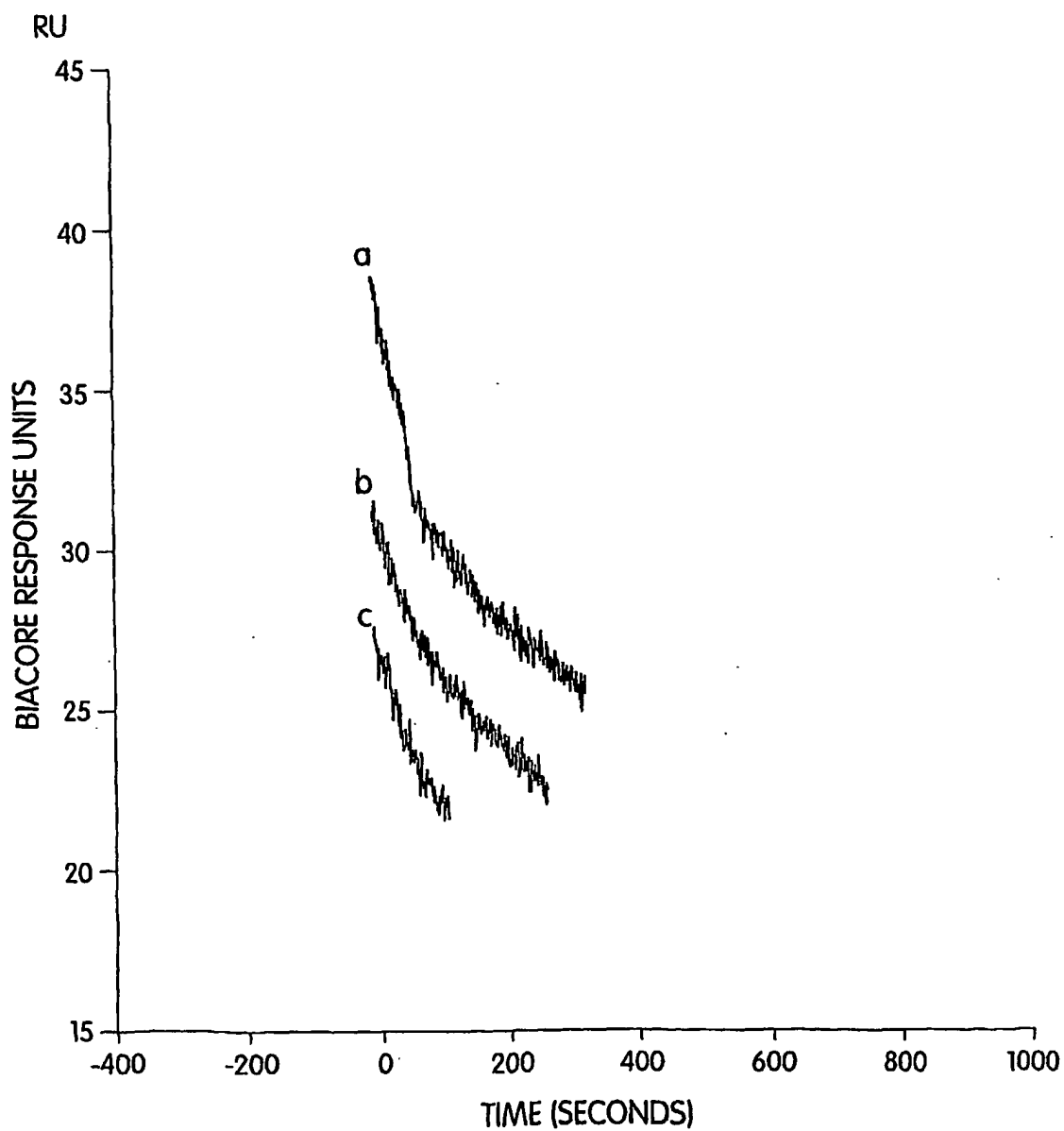


Fig. 10A

**SUBSTITUTE SHEET (RULE 26)**

18/18

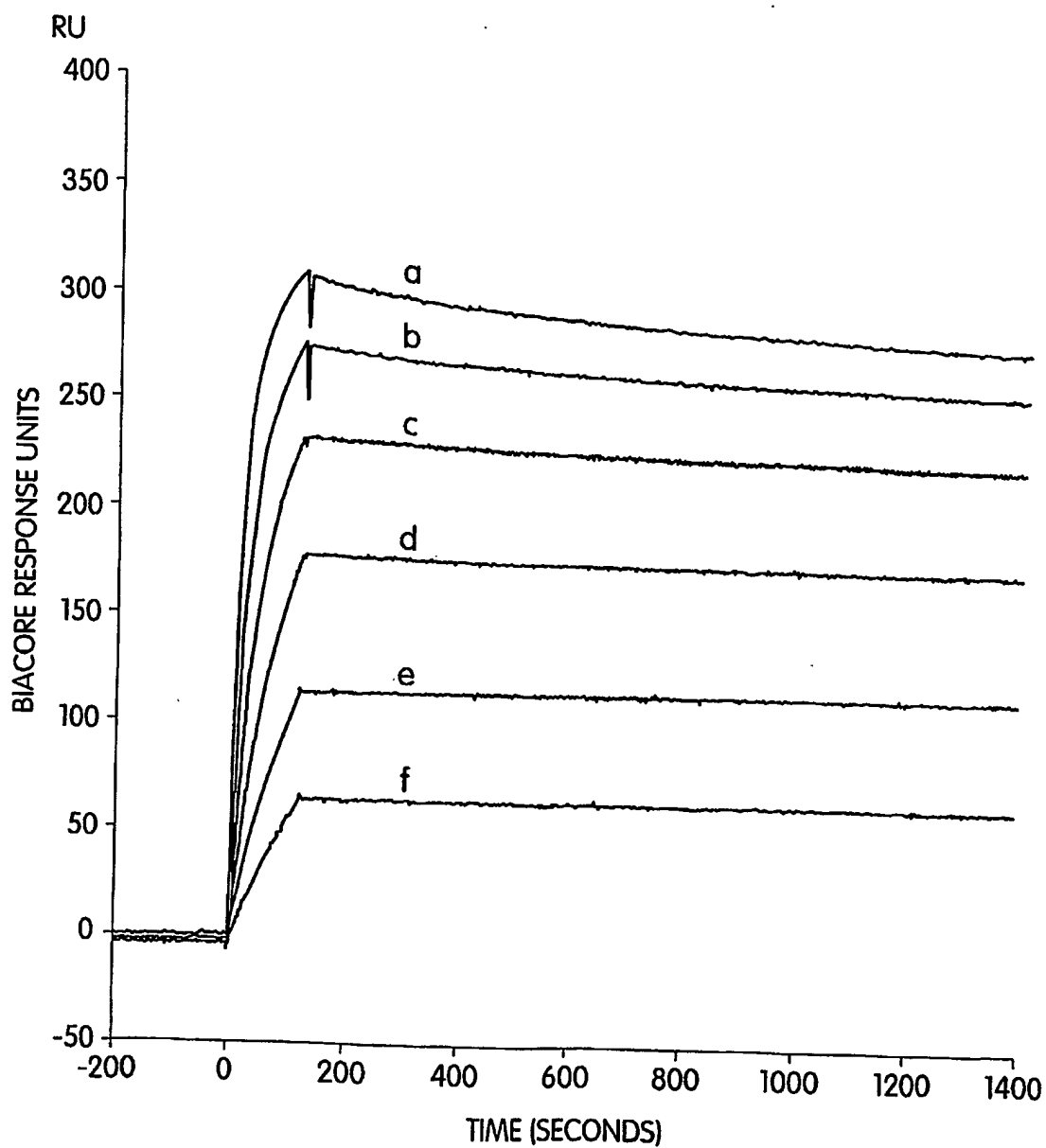


Fig. 10B

**SUBSTITUTE SHEET (RULE 26)**

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/41717

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : Please See Extra Sheet.

US CL : Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC.

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/320.1, 69.1, 4, 7.1, 7.5, 9.92, 69.7; 436/501; 530/300, 412, 344; 536/23.1, 23.4

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
sequence search, STN, EAST, search terms: streptavidin, peptide, bind?, dissociation constant, inventor search.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ----- A	US 6,103,493 A (SKERRA et al) 15 August 2000, entire reference, especially Col. 7, lines 54-55.	1 ----- 2-26
A	US 5,506,121 A (SKERRA et al) 09 April 1996, entire document.	1-26
A	KATZ et al., In Crystals of Complexes of Streptavidin with Peptide Ligands Containing the HPQ Sequence the pKa of the Peptide Histidine is Less than 3.0, J. Biol. Chem., 16 May 1997, Vol. 272, No. 20, pp. 13220-13228, especially p. 13223, Table II.	1-26
A	DEVLIN et al. Random Peptide Libraries: A Source of Specific Protein Binding Molecules, Science, 1990, Vol. 249, pp. 404-406.	1-26

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Z" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

06 SEPTEMBER 2001

Date of mailing of the international search report

19 NOV 2001

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

HOLLY SCHNIZER

Telephone No. (703) 308-0186

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/41717

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	OSTERGAARD et al. Novel Avidin and Streptavidin Binding Sequences Found in Synthetic Peptide Libraries, 1995, Vol. 362, pp. 306-308, especially p. 308, last two paragraphs.	1-26
A	ZANG et al. Tight-Binding Streptavidin Ligands From a Cyclic Peptide Library, 1998, Bioorg. Med. Chem. Lett. Vol. 8, pp. 2327-2332, especially p. 2327, last paragraph.	1-26
A	GIEBEL et al. Screening of Cyclic Peptide Phage Libraries Identifies Ligands That Bind Streptavidin with High Affinities, 1995, Biochemistry, Vol. 34, pp15430-15435, especially p. 15435, Col. 1, lines 1-13.	1-26

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/+1717

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (7):

C07H 21/02, 21/04; C07K 1/00, 5/00; C12N 15/00, 15/63; C12Q 1/00; G01N 33/53

A. CLASSIFICATION OF SUBJECT MATTER:

US CL :

435/320.1, 69.1, 4, 7.1, 7.5, 9.92, 69.7; 436/501; 530/300, 412, 344; 536/23.1, 23.4